

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2005, 11:56:10 ; Search time 113 Seconds

(without alignments)

1419.044 Million cell updates/sec

Title: US-09-879-248-6 *complete*

Perfect score: 2310

Sequence: 1 MSILTLNNTSSPGLFQSG.....LGDVENHYKVPMSANLKVAE 447

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : A\_Geneseq\_23Sep04:\*

1: Geneseqp1980s:\*

2: Geneseqp1990s:\*

3: Geneseqp2000s:\*

4: Geneseqp2001s:\*

5: Geneseqp2002s:\*

6: Geneseqp2003as:\*

7: Geneseqp2003bs:\*

8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2310	100.0	447	2	AAW96260 Hypersens
2	2310	100.0	447	3	AAW71094 Erwinia a
3	2310	100.0	447	3	AAW94855 A hyperse
4	2310	100.0	447	5	AAO22548 Hypersens
5	2310	100.0	447	5	AAE18296 Erwinia a
6	2310	100.0	447	5	AAE16448
7	2310	100.0	447	5	ABO92226 Erwinia a
8	559	24.2	424	2	AAW97851
9	559	24.2	424	3	AAW71098
10	559	24.2	424	3	AAW94859 A hyperse
11	559	24.2	424	5	AAE18238
12	559	24.2	424	5	AAE16452
13	559	24.2	424	5	ABO92228
14	359.5	15.6	197	2	AAW83014 Bacillus
15	359.5	15.6	197	2	AAW7412 Bacillus
16	359.5	15.6	197	2	AAW28446
17	341	14.8	221	2	AAW28446
18	341	14.8	221	8	ADK48880 Bacillus
19	199.5	8.6	1306	6	ABU36481 Protein e
20	193	8.4	694	6	ABU36802 Protein e
21	191.5	8.3	594	6	ABU22680 Protein e
22	191	8.3	2280	4	ABW61650 Drosophil
23	188	8.1	639	6	ABU36685 Protein e
24	186.5	8.1	591	6	ABU36439 Protein e
25	185.5	8.0	1381	6	ABU36971 Protein e

Adm40790	Mycobacte	7	ADM40790	588	8.0	185	26
Abu36945	Protein e	6	ABU36945	588	7.9	183	27
Abu36540	Protein e	6	ABU36540	606	7.9	182	28
Abm15873	Mycobacte	6	ABM15873	484	7.9	181.5	29
Abu36943	Protein e	6	ABU36943	484	7.9	181.5	30
Abu36943	Mycobacte	6	ABU36943	484	7.9	181.5	31
Abu36943	Mycobacte	6	ABU36943	484	7.9	181.5	32
Abu36943	Mycobacte	6	ABU36943	484	7.9	181.5	33
Abu36943	Mycobacte	6	ABU36943	484	7.9	181.5	34
Abu36943	Mycobacte	6	ABU36943	484	7.9	181.5	35
Abu36943	Mycobacte	6	ABU36943	484	7.9	181.5	36
Abu36943	Mycobacte	6	ABU36943	484	7.9	181.5	37
Abu36943	Mycobacte	6	ABU36943	484	7.9	181.5	38
Abu36943	Mycobacte	6	ABU36943	484	7.9	181.5	39
Abu36943	Mycobacte	6	ABU36943	484	7.9	181.5	40
Abu36943	Mycobacte	6	ABU36943	484	7.9	181.5	41
Abu36943	Mycobacte	6	ABU36943	484	7.9	181.5	42
Abu36943	Mycobacte	6	ABU36943	484	7.9	181.5	43
Abu36943	Mycobacte	6	ABU36943	484	7.9	181.5	44
Abu36943	Mycobacte	6	ABU36943	484	7.9	181.5	45
Abu36943	Mycobacte	6	ABU36943	484	7.9	181.5	46
Abu36943	Mycobacte	6	ABU36943	484	7.9	181.5	47
Abu36943	Mycobacte	6	ABU36943	484	7.9	181.5	48
Abu36943	Mycobacte	6	ABU36943	484	7.9	181.5	49
Abu36943	Mycobacte	6	ABU36943	484	7.9	181.5	50
Abu36943	Mycobacte	6	ABU36943	484	7.9	181.5	51
Abu36943	Mycobacte	6	ABU36943	484	7.9	181.5	52
Abu36943	Mycobacte	6	ABU36943	484	7.9	181.5	53
Abu36943	Mycobacte	6	ABU36943	484	7.9	181.5	54
Abu36943	Mycobacte	6	ABU36943	484	7.9	181.5	55
Abu36943	Mycobacte	6	ABU36943	484	7.9	181.5	56
Abu36943	Mycobacte	6	ABU36943	484	7.9	181.5	57
Abu36943	Mycobacte	6	ABU36943	484	7.9	181.5	58
Abu36943	Mycobacte	6	ABU36943	484	7.9	181.5	59
Abu36943	Mycobacte	6	ABU36943	484	7.9	181.5	60
Abu36943	Mycobacte	6	ABU36943	484	7.9	181.5	61
Abu36943	Mycobacte	6	ABU36943	484	7.9	181.5	62
Abu36943	Mycobacte	6	ABU36943	484	7.9	181.5	63
Abu36943	Mycobacte	6	ABU36943	484	7.9	181.5	64
Abu36943	Mycobacte	6	ABU36943	484	7.9	181.5	65
Abu36943	Mycobacte	6	ABU36943	484	7.9	181.5	66
Abu36943	Mycobacte	6	ABU36943	484	7.9	181.5	67
Abu36943	Mycobacte	6	ABU36943	484	7.9	181.5	68
Abu36943	Mycobacte	6	ABU36943	484	7.9	181.5	69
Abu36943	Mycobacte	6	ABU36943	484	7.9	181.5	70
Abu36943	Mycobacte	6	ABU36943	484	7.9	181.5	71
Abu36943	Mycobacte	6	ABU36943	484	7.9	181.5	72
Abu36943	Mycobacte	6	ABU36943	484	7.9	181.5	73
Abu36943	Mycobacte	6	ABU36943	484	7.9	181.5	74
Abu36943	Mycobacte	6	ABU36943	484	7.9	181.5	75
Abu36943	Mycobacte	6	ABU36943	484	7.9	181.5	76
Abu36943	Mycobacte	6	ABU36943	484	7.9	181.5	77
Abu36943	Mycobacte	6	ABU36943	484	7.9	181.5	78
Abu36943	Mycobacte	6	ABU36943	484	7.9	181.5	79
Abu36943	Mycobacte	6	ABU36943	484	7.9	181.5	80
Abu36943	Mycobacte	6	ABU36943	484	7.9	181.5	81
Abu36943	Mycobacte	6	ABU36943	484	7.9	181.5	82
Abu36943	Mycobacte	6	ABU36943	484	7.9	181.5	83
Abu36943	Mycobacte	6	ABU36943	484	7.9	181.5	84
Abu36943	Mycobacte	6	ABU36943	484	7.9	181.5	85
Abu36943	Mycobacte	6	ABU36943	484	7.9	181.5	86
Abu36943	Mycobacte	6	ABU36943	484	7.9	181.5	87
Abu36943	Mycobacte	6	ABU36943	484	7.9	181.5	88
Abu36943	Mycobacte	6	ABU36943	484	7.9	181.5	89
Abu36943	Mycobacte	6	ABU36943	484	7.9	181.5	90
Abu36943	Mycobacte	6	ABU36943	484	7.9	181.5	91
Abu36943	Mycobacte	6	ABU36943	484	7.9	181.5	92
Abu36943	Mycobacte	6	ABU36943	484	7.9	181.5	93
Abu36943	Mycobacte	6	ABU36943	484	7.9	181.5	94
Abu36943	Mycobacte	6	ABU36943	484	7.9	181.5	95
Abu36943	Mycobacte	6	ABU36943	484	7.9	181.5	96
Abu36943	Mycobacte	6	ABU36943	484	7.9	181.5	97
Abu36943	Mycobacte	6	ABU36943	484	7.9	181.5	98



QY 121 QGGGQIGDNPFLKAMLLKLIARMDGSDQFGQPGTGNNSASSGTSSGGSPFNDSGGKA 180  
DB 121 QGGGQIGDNPFLKAMLLKLIARMDGSDQFGQPGTGNNSASSGTSSGGSPFNDSGGKA 180  
QY 181 PSNGSPSGNYSVSTSPSTPTSPPLDPSSPTKAGGSTPVTDHDPVGSAGIGAG 240  
DB 181 PSNGSPSGNYSVSTSPSTPTSPPLDPSSPTKAGGSTPVTDHDPVGSAGIGAG 240  
QY 241 NSVAFTSAGANQTVLHDTITVKAGQVDFGKGQFTTAGSELGQSGQSENQKPLFILEDGAS 300  
DB 241 NSVAFTSAGANQTVLHDTITVKAGQVDFGKGQFTTAGSELGQSGQSENQKPLFILEDGAS 300  
QY 301 LKNVTMGDDGADGHIHYGDGAKIDNLHVTNVGDAITVKPNSAGKSHVEITNSPFEHSD 360  
DB 301 LKNVTMGDDGADGHIHYGDGAKIDNLHVTNVGDAITVKPNSAGKSHVEITNSPFEHSD 360  
QY 361 KILQLNADTNLSVDNVKAKDFGTFVRTNGGQGNWDLNLSHISAEDGKFSFVKSDSEGLN 420  
DB 361 KILQLNADTNLSVDNVKAKDFGTFVRTNGGQGNWDLNLSHISAEDGKFSFVKSDSEGLN 420  
QY 421 VNTSDISLGDVENHYKVPMSANLKVAE 447  
DB 421 VNTSDISLGDVENHYKVPMSANLKVAE 447  
RESULT 3  
ID AAY84855  
XX AAY84855 standard; protein; 447 AA.  
AC AAY84855;  
XX  
DT 08-AUG-2000 (first entry)  
DE A hypersensitive response elicitor protein.  
XX  
KW Hypersensitive response; insect control; disease resistance;  
KW Hypersensitive response elicitor; plant growth; vegetable; crop;  
KW ornamental plant.  
XX  
OS Erwinia amylovora.  
XX  
PN WO200020452-A2.  
XX  
PD 13-APR-2000.  
XX  
PF 05-OCT-1999; 99WO-US023181.  
XX  
PR 05-OCT-1998; 98US-0103050P.  
XX  
PA (EDEN-) EDEN BIOSCIENCE CORP.  
XX  
PI Wei Z, Fan H, Niggemeyer JL;  
XX  
XX WPI; 2000-303745/26.  
DR N-PSDB; AAA14939.  
XX  
PT Hypersensitive response elicitor polypeptides useful for imparting  
PT enhanced growth, disease resistance and insect resistance to plants,  
PT especially vegetables and ornamental flowers.  
XX  
PS Disclosure; Page 12-13; 100pp; English.  
XX  
CC The present sequence represents a hypersensitive response elicitor  
CC polypeptide. The specification describes hypersensitive response elicitor  
CC polypeptide fragments, which do not elicit a hypersensitive response.  
CC Instead, the proteins impart disease resistance to plants, enhance plant  
CC growth, and/or control insects. The polypeptide fragments may be used to  
CC these properties to plants. The plants which may be treated in this way  
CC include vegetables, crops and ornamental plants such as alfalfa, rice,  
CC wheat, barley, rye, cotton, sunflower, peanut, corn, potato, sweet  
CC potato, bean, pea, chickory, lettuce, endive, cabbage, brussel sprout,  
CC beet, parsnip, turnip, cauliflower, broccoli, radish, spinach, onion,

CC garlic, eggplant, pepper, celery, carrot, squash, pumpkin, zucchini,  
CC cucumber, apple, pear, melon, citrus, strawberry, grape, raspberry,  
CC pineapple, soybean, tobacco, tomato, sugarcane, Arabidopsis  
CC thaliana, Saintpaulia, petunia, pelargonium, poinsettia, chrysanthemum,  
CC carnation or zinnia  
XX  
SQ Sequence 447 AA;  
Query Match 100.0%; Score 2310; DB 3; Length 447;  
Best Local Similarity 100.0%; Pred. No. 2.4e-151;  
Matches 447; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MSILTLNNNTSSPGLFQSGDNGLGGHNANSALGOOPIDRQTIQMAQLLAELLKSLLS 60  
DB 1 MSILTLNNNTSSPGLFQSGDNGLGGHNANSALGOOPIDRQTIQMAQLLAELLKSLLS 60  
QY 61 POSGNAATCAGGNDQTTGVGNAGLNGRGTAGTTTQSDSQNNMLSEWNGNGLDQAITPDG 120  
DB 61 POSGNAATCAGGNDQTTGVGNAGLNGRGTAGTTTQSDSQNNMLSEWNGNGLDQAITPDG 120  
QY 121 QGGGQIGDNPFLKAMLLKLIARMDGSDQFGQPGTGNNSASSGTSSGGSPFNDSGGKA 180  
DB 121 QGGGQIGDNPFLKAMLLKLIARMDGSDQFGQPGTGNNSASSGTSSGGSPFNDSGGKA 180  
QY 181 PSNGSPSGNYSVSTSPSTPTSPPLDPSSPTKAGGSTPVTDHDPVGSAGIGAG 240  
DB 181 PSNGSPSGNYSVSTSPSTPTSPPLDPSSPTKAGGSTPVTDHDPVGSAGIGAG 240  
QY 241 NSVAFTSAGANQTVLHDTITVKAGQVDFGKGQFTTAGSELGQSGQSENQKPLFILEDGAS 300  
DB 241 NSVAFTSAGANQTVLHDTITVKAGQVDFGKGQFTTAGSELGQSGQSENQKPLFILEDGAS 300  
QY 301 LKNVTMGDDGADGHIHYGDGAKIDNLHVTNVGDAITVKPNSAGKSHVEITNSPFEHSD 360  
DB 301 LKNVTMGDDGADGHIHYGDGAKIDNLHVTNVGDAITVKPNSAGKSHVEITNSPFEHSD 360  
QY 361 KILQLNADTNLSVDNVKAKDFGTFVRTNGGQGNWDLNLSHISAEDGKFSFVKSDSEGLN 420  
DB 361 KILQLNADTNLSVDNVKAKDFGTFVRTNGGQGNWDLNLSHISAEDGKFSFVKSDSEGLN 420  
QY 421 VNTSDISLGDVENHYKVPMSANLKVAE 447  
DB 421 VNTSDISLGDVENHYKVPMSANLKVAE 447  
RESULT 4  
ID AAO22548  
XX AAO22548 standard; protein; 447 AA.  
AC AAO22548;  
XX  
DT 28-OCT-2002 (first entry)  
DE Hypersensitive response elicitor protein #2 from Erwinia amylovora.  
XX  
KW Desiccation inhibitor; longevity enhancer; desiccation inhibitor;  
KW cutting; ornamental plant; hypersensitive response elicitor protein;  
KW transgenic ornamental plant; early flowering; longevity enhancer; plant;  
KW flower bloom; flower.  
XX  
OS Erwinia amylovora.  
XX  
PN WO200237960-A2.  
XX  
PD 16-MAY-2002.  
XX  
PF 06-NOV-2001; 2001WO-US043715.  
XX  
PR 13-NOV-2000; 2000US-0248169P.  
XX  
PA (EDEN-) EDEN BIOSCIENCE CORP.  
XX  
PI Wei Z, Leon B, Oviedo A;

```
XX WPI; 2002-575194/61.
DR N-PSDB; AAL41133.
XX
PT Inhibiting desiccation of cuttings from ornamental plants, by treating
PT ornamental plants with hypersensitive response elicitor protein, or
PT expressing heterologous hypersensitive response elicitor protein in
PT plants.
XX
PS Disclosure; Page 12-13; 69pp; English.
XX
CC The invention relates to a method for inhibiting desiccation of cuttings
CC from ornamental plants. The method involves treating the cuttings with a
CC hypersensitive response elicitor protein or polypeptide, or providing a
CC transgenic ornamental plant or plant seed transformed with a DNA molecule
CC encoding the hypersensitive response elicitor polypeptide, and growing
CC the ornamental plant or transgenic ornamental plant produced from the
CC transgenic ornamental plant seed. The hypersensitive response elicitor
CC protein or polypeptide is useful for inhibiting desiccation of cuttings
CC from ornamental plants, for harvesting cuttings from ornamental plants,
CC for promoting early flowering of ornamental plants, and enhancing the
CC longevity of flower blooms on ornamental plant cuttings. This sequence
CC represents a hypersensitive response elicitor protein of the invention
XX
SQ Sequence 447 AA;

Query Match      100.0%; Score 2310; DB 5; Length 447;
Best Local Similarity 100.0%; Pred. No. 2.4e-151;
Matches 447; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSILTLNNTSSPGLFQSGDNGLGCHNANSALGQPIDRQTTEQMAQLLAELLKSLLS 60
Db 1 MSILTLNNTSSPGLFQSGDNGLGCHNANSALGQPIDRQTTEQMAQLLAELLKSLLS 60
QY 61 POSGNAATGAGNDQTTGVGNAGGLNGRKGTTAGTTPQSDSQNMLSEMGNGLDQAITPDG 120
Db 61 POSGNAATGAGNDQTTGVGNAGGLNGRKGTTAGTTPQSDSQNMLSEMGNGLDQAITPDG 120
QY 121 QGGGQIGDNPLLKAMLLIARMDGQSDQFGQPGTGNNNSASSGTSSSGSPFNDLSGGKA 180
Db 121 QGGGQIGDNPLLKAMLLIARMDGQSDQFGQPGTGNNNSASSGTSSSGSPFNDLSGGKA 180
QY 181 PSGNSPSGNYSPVSTFSPSTPTSPDPPSPPTKAAGGSTPTVTHDPVPVGSAGIGAG 240
Db 181 PSGNSPSGNYSPVSTFSPSTPTSPDPPSPPTKAAGGSTPTVTHDPVPVGSAGIGAG 240
QY 241 NSVAFTSAGANQTVLHDTITVKAGQVFDGKGQTTAGSELGQSQENQKPLFILEDGAS 300
Db 241 NSVAFTSAGANQTVLHDTITVKAGQVFDGKGQTTAGSELGQSQENQKPLFILEDGAS 300
QY 301 LKNVTMGDDGADGHIHYGDAKIDNLHVTNVGDEAITVKPNSAGKSHVEITNSSFEHSD 360
Db 301 LKNVTMGDDGADGHIHYGDAKIDNLHVTNVGDEAITVKPNSAGKSHVEITNSSFEHSD 360
QY 361 KILQLNADTNLSVDNVKAKDFGTFVRTNGGQGNWDLNLSHISAEDGKFSFVKSDSEGLN 420
Db 361 KILQLNADTNLSVDNVKAKDFGTFVRTNGGQGNWDLNLSHISAEDGKFSFVKSDSEGLN 420

RESULT 5
AAE18296
ID AAE18296 standard; protein; 447 AA.
XX
AC AAE18296;
XX
DT 07-MAY-2002 (first entry)
XX
DE Erwinia amylovora hypersensitive response elicitor (HRE) #2.
XX
```

```
KW Hypersensitive response elicitor; HRE; transgenic plant; plant growth;
KW stress tolerance; disease tolerance; modified flower colour;
KW insect resistance; herbicide resistance; male sterility.
XX
OS Erwinia amylovora.
XX WO200195724-A2.
XX
PN 20-DEC-2001.
XX
PD 13-JUN-2001; 2001WO-US018955.
XX
PF 15-JUN-2000; 2000US-0211585P.
XX
PR (EDEN-) EDEN BIOSCIENCE CORP.
XX
PA Wei Z, Derocher J;
XX WPI; 2002-130707/17.
XX N-PSDB; AAD29125.
XX
PT Improving effectiveness of transgenic plants by topical application of a
PT hypersensitive response elicitor protein to the transgenic plant or by
PT incorporating into the plant a transgene encoding the protein.
XX
PS Disclosure; Page 13-14; 86pp; English.
XX
CC The invention relates to methods of improving the effectiveness of
CC transgenic plants which involves either topical application of a
CC hypersensitive response elicitor (HRE) protein to the transgenic plant or
CC incorporating into the transgenic plant a transgene encoding HRE. HRE
CC sequence is used for improving the effectiveness of transgenic plants by
CC maximising the benefit of transgenic traits associated with a deleterious
CC effect on growth, stress tolerance, disease or insect resistance,
CC enhanced growth, herbicide resistance, male sterility, modified flower
CC colour and biochemically modified plant product in the transgenic plants
CC or overcoming the deleterious effects. The present sequence is Erwinia
CC amylovora HRE protein
XX
SQ Sequence 447 AA;

Query Match      100.0%; Score 2310; DB 5; Length 447;
Best Local Similarity 100.0%; Pred. No. 2.4e-151;
Matches 447; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSILTLNNTSSPGLFQSGDNGLGCHNANSALGQPIDRQTTEQMAQLLAELLKSLLS 60
Db 1 MSILTLNNTSSPGLFQSGDNGLGCHNANSALGQPIDRQTTEQMAQLLAELLKSLLS 60
QY 61 POSGNAATGAGNDQTTGVGNAGGLNGRKGTTAGTTPQSDSQNMLSEMGNGLDQAITPDG 120
Db 61 POSGNAATGAGNDQTTGVGNAGGLNGRKGTTAGTTPQSDSQNMLSEMGNGLDQAITPDG 120
QY 121 QGGGQIGDNPLLKAMLLIARMDGQSDQFGQPGTGNNNSASSGTSSSGSPFNDLSGGKA 180
Db 121 QGGGQIGDNPLLKAMLLIARMDGQSDQFGQPGTGNNNSASSGTSSSGSPFNDLSGGKA 180
QY 181 PSGNSPSGNYSPVSTFSPSTPTSPDPPSPPTKAAGGSTPTVTHDPVPVGSAGIGAG 240
Db 181 PSGNSPSGNYSPVSTFSPSTPTSPDPPSPPTKAAGGSTPTVTHDPVPVGSAGIGAG 240
QY 241 NSVAFTSAGANQTVLHDTITVKAGQVFDGKGQTTAGSELGQSQENQKPLFILEDGAS 300
Db 241 NSVAFTSAGANQTVLHDTITVKAGQVFDGKGQTTAGSELGQSQENQKPLFILEDGAS 300
QY 301 LKNVTMGDDGADGHIHYGDAKIDNLHVTNVGDEAITVKPNSAGKSHVEITNSSFEHSD 360
Db 301 LKNVTMGDDGADGHIHYGDAKIDNLHVTNVGDEAITVKPNSAGKSHVEITNSSFEHSD 360
QY 361 KILQLNADTNLSVDNVKAKDFGTFVRTNGGQGNWDLNLSHISAEDGKFSFVKSDSEGLN 420
Db 361 KILQLNADTNLSVDNVKAKDFGTFVRTNGGQGNWDLNLSHISAEDGKFSFVKSDSEGLN 420
```



QY 421 VNTSDISLGDVENHYKVPMSANLKVAE 447  
Db 421 VNTSDISLGDVENHYKVPMSANLKVAE 447

RESULT 6  
AAE16448  
ID AAE16448 standard; protein; 447 AA.  
XX AC AAE16448;  
XX 09-APR-2002 (first entry)  
XX DE E. amylovora hypersensitive response elicitor protein #2.  
XX KW Hypersensitive response elicitor protein; plant growth; fruit coloration;  
KW disease resistance; stress resistance; phytotoxin; insect infection;  
KW plant maturation.  
XX OS Erwinia amylovora.  
XX FH Key Location/Qualifiers  
FT Domain 5..64  
FT FT /label= Hypersensitive\_response\_eliciting\_domain\_1  
FT Region 5..45  
FT FT /label= Acidic\_unit  
FT Region 45..64  
FT FT /label= Alpha\_helix  
FT Domain 103..146  
FT FT /label= Hypersensitive\_response\_eliciting\_domain\_2  
FT Region 103..131  
FT FT /label= Acidic\_unit  
FT Region 131..146  
FT FT /label= Alpha\_helix  
XX WO200198501-A2.  
XX 27-DEC-2001.  
XX 12-JUN-2001; 2001WO-US018820.  
XX 16-JUN-2000; 2000US-0212211P.  
XX (EDEN-) EDEN BIOSCIENCE CORP.  
XX Fan H, Wei Z;  
XX WPI; 2002-122282/16.  
XX N-PSDB; AAD27016.  
XX New hypersensitive response elicitor proteins comprising spaced apart  
PT domains having an acidic portion linked to an alpha-helix, useful for  
PT imparting disease or stress resistance, controlling insects or enhancing  
PT plant growth.  
XX Disclosure; Page 13-14; 99pp; English.  
XX The patent discloses hypersensitive response elicitor proteins and  
CC nucleotides encoding such proteins. Hypersensitive response elicitor  
CC proteins comprise an isolated pair or more of spaced apart domains, each  
CC comprising an acidic portion linked to an alpha-helix and capable of  
CC eliciting a hypersensitive response in plants. Sequences of the invention  
CC are used to impart disease resistance to plants, to enhance plant growth,  
CC to control insects and/or to impart stress resistance to plants which  
CC includes resistance to environmental stresses such as climate, air  
CC pollution, chemical and nutritional stress. The method of imparting  
CC disease resistance has the potential for treating previously untreatable  
CC diseases, treating diseases systemically and avoiding the use of  
CC infectious agents or environmentally harmful materials. Hyper-sensitive  
CC response elicitor sequences are used to enhance plant growth which  
CC encompasses greater yield, increased in quantity of seeds produced,  
CC percentage of seeds germinated, plant size and biomass, bigger fruits,  
CC earlier fruit coloration and plant maturation. They are also used for

CC insect control which encompasses preventing direct insect damage to plant  
CC by feeding injury, interfering with insect larval feeding on the plants,  
CC preventing insects from colonising host plants and releasing phytotoxins.  
CC Sequences of the invention also prevent subsequent disease damage to  
CC plants resulting from insect infection. The present sequence is Erwinia  
CC amylovora hypersensitive response elicitor protein  
XX SQ Sequence 447 AA;  
Query Match 100.0%; Score 2310; DB 5; Length 447;  
Best Local Similarity 100.0%; Pred. No. 2.4e-151;  
Matches 447; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MSILTLNNTSSPGLFQSGDNLGCHNANSALGOQPIDRQTIEQWALAEKLSLLS 60  
Db 1 MSILTLNNTSSPGLFQSGDNLGCHNANSALGOQPIDRQTIEQWALAEKLSLLS 60  
QY 61 PQSGNAATCAGGNDQTTGVGNAGLNGRGTAGTTTQSDSQNNMLSEWNNGLDQAITPDG 120  
Db 61 PQSGNAATCAGGNDQTTGVGNAGLNGRGTAGTTTQSDSQNNMLSEWNNGLDQAITPDG 120  
QY 121 QGGGQIGDNPLKAMLKLIARMMDGSDQFGQPGTGNNSSASCTSSSGSPFNDLSGKA 180  
Db 121 QGGGQIGDNPLKAMLKLIARMMDGSDQFGQPGTGNNSSASCTSSSGSPFNDLSGKA 180  
QY 181 PSGNSPSGNYSPVSTPSTPTSPDLPSSPTKAAAGGSTPTVTDHPDVPVGSAGIGAG 240  
Db 181 PSGNSPSGNYSPVSTPSTPTSPDLPSSPTKAAAGGSTPTVTDHPDVPVGSAGIGAG 240  
QY 241 NSVAFTSAGANQTVLHDTITVKAGQVFDGKGTFTAGSELGDSQENQKPLFIEGDAS 300  
Db 241 NSVAFTSAGANQTVLHDTITVKAGQVFDGKGTFTAGSELGDSQENQKPLFIEGDAS 300  
QY 301 LKNVTMGDDGADGIHLYGDAKIDNLHVTNVGDAITVKPNSACKKSHVEITNSFEHASD 360  
Db 301 LKNVTMGDDGADGIHLYGDAKIDNLHVTNVGDAITVKPNSACKKSHVEITNSFEHASD 360  
QY 361 KILQLNADTNLSVDNVKAKDFGFTVTNGGQQNQNDNLNLSHISAEDGKFSFVKSDSEGLN 420  
Db 361 KILQLNADTNLSVDNVKAKDFGFTVTNGGQQNQNDNLNLSHISAEDGKFSFVKSDSEGLN 420  
QY 421 VNTSDISLGDVENHYKVPMSANLKVAE 447  
Db 421 VNTSDISLGDVENHYKVPMSANLKVAE 447

RESULT 7  
ABB09226  
ID ABB09226 standard; protein; 447 AA.  
XX AC ABB09226;  
XX 08-JUL-2002 (first entry)  
XX DE Erwinia amylovora hypersensitive response elicitor SEQ ID NO:5.  
XX KW Hypersensitive response elicitor; fruit; vegetable; plant; desiccation;  
XX postharvest disease.  
XX OS Erwinia amylovora.  
XX WO200180639-A2.  
XX 01-NOV-2001.  
XX 17-APR-2001; 2001WO-US012468.  
XX 19-APR-2000; 2000US-0198359P.  
XX (EDEN-) EDEN BIOSCIENCE CORP.  
XX Wei Z, Qiu D, Remick D;  
XX



Db 201 NPNRTSG-----TGSVSDTAGS-----TEQAGKINNVKDTIKVGAGEVFDGHGAT 245

Qy 274 FTAGSELGCGQSENOKPLFILEDGASLKNVTMGDDGADGIHLYG-----DAKIDNHLVTN 329

Db 246 FTADKSMGNDQGENOKPMFELAEAGATLKNVNLGENEVDGIHVKAQNAQEVTTIDNVHAQN 305

Qy 330 VGEDAITVKPNSAGKSHVEITNSSPEHASDKILOLNADTNLSVDNVKAKDFGTFVVRTNG 389

Db 306 VGEDLITVKEGGAAVTNLNKNSSAKGADDKVQLNANHLKIDNFKADDFGTFVVRTNG 365

Qy 390 GQOQ-GNWDNLNLSHISAEADGKFSFKVSDSEGLNVTNTSDISLGDYENHY 435

Db 366 GKQFDDMSIELNGIEANHGKFAVLSKSDSDLLKATGNIAMTDVKHAY 412

RESULT 9

AAV71098

ID AAY71098 standard; protein; 424 AA.

XX AC AAY71098;

DT 08-SEP-2000 (first entry)

XX Pseudomonas syringae dspE gene encoded hypersensitive response elicitor.

XX Hypersensitive response elicitor; environmental stress resistance; plant;

KW pathogen; dspE gene.

XX Pseudomonas syringae.

OS Pseudomonas syringae.

XX WO200028055-A2.

PN 18-MAY-2000.

PD 04-NOV-1999; 99WO-US026039.

PF 05-NOV-1998; 98US-010743P.

PR (EDEN-) EDEN BIOSCIENCE CORP.

XX Wei Z, Schading RL;

PI WPI; 2000-376566/32.

XX N-PSDB; AAD00673.

DR Application of a hypersensitive response elicitor protein to plants to

PT impart stress resistance.

XX Disclosure; Page 25-26; 84pp; English.

XX The patent discloses a method to impart stress resistance to plants by

CC applying a hypersensitive response elicitor in a non-infectious form to a

CC plant or seed. The present sequence is a hypersensitive response elicitor

CC protein encoded by dspE gene from Pseudomonas syringae. The protein

CC elicits a plant pathogen's hypersensitive response and is used to impart

CC stress resistance to plants

XX Sequence 424 AA;

SQ

Query Match 24.2%; Score 559; DB 3; Length 424;

Best Local Similarity 36.1%; Pred. No. 2.7e-30;

Matches 147; Conservative 54; Mismatches 132; Indels 74; Gaps 13;

Qy 37 QPIDRTIETQMAQLLAELKSL---LSPQSGNAATGAGGNDQTTGVNAGLNGRKGTAG 93

Db 72 KPNDSSG--NIAKLISALINSLQLMLTNSNKKQDTNQEQDPSQAPQNNGLG-----122

Qy 94 TTPQSDSQNMLSBMGNGLDQAITPDGQGGGQIGDNPFLKAMKLKLIARMMDGSDQFGQP 153

Db 123 -TPSADS-----GGGG-----TPDATGGG-GDTP-----SATGGG 151

Qy 154 GTGNNGASSGTSSTSGSPFNLDLGGKAPSGNSPVSFTSPSTPTSPSTPLDFFPS 213

Db 152 GGDTPATATGGGGGGTPTATGG---SGGTPATATGGGGVTPQITPOL-----A 200

Qy 214 SPTKAAAGSTPVTYDHPDPVGSAGIGAGNSVAFVTSAGANQTVLHDTTIVKAGQVFDKGGOT 273

Db 201 NPNRTSG-----TGSVSDTAGS-----TEQAGKINNVKDTIKVGAGEVFDGHGAT 245

Qy 274 FTAGSELGCGQSENOKPLFILEDGASLKNVTMGDDGADGIHLYG-----DAKIDNHLVTN 329

Db 246 FTADKSMGNDQGENOKPMFELAEAGATLKNVNLGENEVDGIHVKAQNAQEVTTIDNVHAQN 305

Qy 330 VGEDAITVKPNSAGKSHVEITNSSPEHASDKILOLNADTNLSVDNVKAKDFGTFVVRTNG 389

Db 306 VGEDLITVKEGGAAVTNLNKNSSAKGADDKVQLNANHLKIDNFKADDFGTFVVRTNG 365

Qy 390 GQOQ-GNWDNLNLSHISAEADGKFSFKVSDSEGLNVTNTSDISLGDYENHY 435

Db 366 GKQFDDMSIELNGIEANHGKFAVLSKSDSDLLKATGNIAMTDVKHAY 412

RESULT 10

AAV84859

ID AAY84859 standard; protein; 424 AA.

XX AC AAY84859;

DT 08-AUG-2000 (first entry)

XX A hypersensitive response elicitor protein.

XX Hypersensitive response; insect control; disease resistance;

KW hypersensitive response elicitor; plant growth; vegetable; crop;

XX ornamental plant.

OS Pseudomonas syringae.

XX WO200020452-A2.

PN 13-APR-2000.

PD 05-OCT-1999; 99WO-US023181.

PF 05-OCT-1998; 98US-0103050P.

PR (EDEN-) EDEN BIOSCIENCE CORP.

XX Wei Z, Fan H, Niggenmeyer JL;

PI WPI; 2000-303745/26.

XX N-PSDB; AAAL4943.

DR Hypersensitive response elicitor polypeptides useful for imparting

PT enhanced growth, disease resistance and insect resistance to plants,

PT especially vegetables and ornamental flowers.

XX Disclosure; Page 26-28; 100pp; English.

XX The present sequence represents a hypersensitive response elicitor

CC polypeptide. The specification describes hypersensitive response elicitor

CC polypeptide fragments, which do not elicit a hypersensitive response.

CC Instead, the proteins impart disease resistance to plants, enhance plant

CC growth, and/or control insects. The polypeptide fragments may be used to

CC include vegetables, crops and ornamental plants such as alfalfa, rice,

CC wheat, barley, rye, cotton, sunflower, peanut, corn, potato, sweet

CC potato, bean, pea, chickory, lettuce, endive, cabbage, brussel sprout,

CC beet, parsnip, turnip, cauliflower, broccoli, radish, spinach, onion,

CC garlic, eggplant, pepper, celery, carrot, squash, pumpkin, zucchini,

CC cucumber, apple, pear, melon, citrus, strawberry, grape, raspberry,

CC pineapple, soybean, tobacco, tomato, sorghum or sugarcane, Arabidopsis

CC thaliana, Saintpaulia, petunia, pelargonium, poinsettia, chrysanthemum,

CC carnation or zinnia

XX Sequence 424 AA;

SQ







CC protopectinase having an optimum pH 7 or higher against protopectin and  
CC polygalacturonic acid substrates. Inclusion of protopectinase gives a  
CC composition that provides better removal of muddy soil, particularly from  
CC socks. The present sequence represents pectic acid lyase from *Bacillus*  
CC sp. strain KSM-P15, which is used in an example from the present  
CC invention. Pectic acid lyase exhibits protopectinase activity  
SQ Sequence 197 AA;

Query Match 15.6%; Score 359.5; DB 2; Length 197;  
Best Local Similarity 46.3%; Pred. No. 6.2e-17;  
Matches 81; Conservative 29; Mismatches 56; Indels 9; Gaps 4;  
QY 253 TVLHDTITVKGQVFGKGTFTAG-SELGDGQSQENQKPLFILEDGASLKNVTMGDDGA 311  
Db 3 TVVHETIRVPAGQTFDGKGTYYVANTLGDGSAENQKPIFLEAGASLKNVIGAPAA 62  
QY 312 DGIHLYGDAKIDNLHVTNVEDAITVKPNSAGKSKSHVEITNSSFEHASDKILOLNADTNL 371  
Db 63 DGVCYGDCTITNVIWEDVGEDALTUK--SSGT---VNISGGAAYKAYDKVQFQINAAGTI 117  
QY 372 SVDNVKAKDFGTFRVTNGGQGNWDLNLSHISAEDGKFSFVKSDSE---GLNVNT 423  
Db 118 NIRNFRADDIGKLVQRNGGTTYKVMNVENCNISRVKDAILRTDSTSTGRIVNT 172

RESULT 16  
AAW77412  
ID AAW77412 standard; protein; 197 AA.

XX AAW77412;  
AC AAW77412;  
XX 07-JAN-1999 (first entry)  
DT 07-JAN-1999 (first entry)  
XX *Bacillus* sp. pectic acid lyase.  
XX *Bacillus* sp. KSM-P15; pectic acid lyase; pectinic acid lyase; pectin;  
KW detergent; food-processing; fibre-processing agent; cell wall.  
XX *Bacillus* sp.  
OS EP870834-A1.  
PN 14-OCT-1998.  
PD 09-APR-1998; 98EP-00106586.  
PF 09-APR-1997; 97JP-00091142.  
PR 08-SEP-1997; 97JP-00242735.  
XX (KAOS ) KAO CORP.  
XX Hatada Y, Koike K, Yoshimatsu T, Suzumatsu A, Kobayashi T, Ito S;  
PI WPI; 1998-523159/45.  
DR N-PSDB; AAV59478.

XX New *Bacillus* pectic acid lyase - useful as a detergent component, a food-  
PT processing agent and a fibre-processing agent.  
XX Claim 1; Page 16-17; 29pp; English.  
XX The present sequence represents a pectic acid lyase isolated from  
CC microorganism *Bacillus* sp. KSM-P15. The pectic acid lyase has high pectic  
CC acid lyase activity which degrades pectin in plant cell walls and fibre  
CC in vegetables, and so is useful as a component of detergents, a food-  
CC processing agent, or a fibre-processing agent. The pectic acid lyase has  
CC a higher optimum reaction pH (10.3-10.7) than known *Bacillus* pectic acid  
CC lyases (pH 8-9.5) and so has wider industrial applications. Unlike  
CC present pectic acid lyases, the new enzyme has a high enzyme activity,  
CC and can be produced on a mass scale  
SQ Sequence 197 AA;

Query Match 15.6%; Score 359.5; DB 2; Length 197;  
Best Local Similarity 46.3%; Pred. No. 6.2e-17;  
Matches 81; Conservative 29; Mismatches 56; Indels 9; Gaps 4;  
QY 253 TVLHDTITVKGQVFGKGTFTAG-SELGDGQSQENQKPLFILEDGASLKNVTMGDDGA 311  
Db 3 TVVHETIRVPAGQTFDGKGTYYVANTLGDGSAENQKPIFLEAGASLKNVIGAPAA 62  
QY 312 DGIHLYGDAKIDNLHVTNVEDAITVKPNSAGKSKSHVEITNSSFEHASDKILOLNADTNL 371  
Db 63 DGVCYGDCTITNVIWEDVGEDALTUK--SSGT---VNISGGAAYKAYDKVQFQINAAGTI 117  
QY 372 SVDNVKAKDFGTFRVTNGGQGNWDLNLSHISAEDGKFSFVKSDSE---GLNVNT 423  
Db 118 NIRNFRADDIGKLVQRNGGTTYKVMNVENCNISRVKDAILRTDSTSTGRIVNT 172

RESULT 17  
AAZ28446  
ID AAZ28446 standard; protein; 221 AA.

XX AAZ28446;  
XX 17-JAN-2000 (first entry)  
DT 17-JAN-2000 (first entry)  
XX *Bacillus* licheniformis Pectate lyase I.  
DE Pectate lyase I; EC 4.2.2.2; pectin degrading enzyme; ATCC 14580;  
KW catalytic active domain; cellulose binding domain; CBD; operably linked;  
KW optimum activity; pH; detergent composition; yarn; cellulosic fibre;  
KW recycled waste paper; pulp; retting process; animal feed; wine; juice;  
KW transgenic plant.  
XX *Bacillus* licheniformis.  
XX Key Location/Qualifiers  
FH Peptide 1..27 /label= Pro-sequence  
FT Protein 28..221 /label= Mature\_pectate\_lyase\_I  
FT Misc-difference 133 /note= "Conserved residue"  
FT Misc-difference 155 /note= "Conserved residue"

XX WO9927083-A1.  
XX 03-JUN-1999.  
XX 24-NOV-1998; 98WO-DK000514.  
XX 24-NOV-1997; 97DK-00001344.  
PR 06-MAY-1998; 98US-00073684.  
XX (NOVO ) NOVO-NORDISK AS.  
XX Andersen LN, Schuelein M, Lange NEK, Bjornvad ME, Schnorr K;  
PI WPI; 1999-610578/52.  
DR N-PSDB; AAZ89484.  
XX New isolated pectate lyase enzymes.  
XX Claim 6; Page 79-80; 93pp; English.  
XX The present sequence is pectate lyase I, which is a pectin degrading  
CC enzyme derived from *Bacillus* licheniformis, ATCC 14580. The enzyme  
CC comprises a catalytic active domain and a cellulose binding domain (CBD)  
CC that are operably linked to each other. It shows optimum activity at pH  
CC greater than 9 and temperature 55 degree centigrade. It can be used in  
CC detergent compositions, for cleaning hard surfaces, for machine treatment  
CC of fabrics, for improving the properties of cellulosic fibres, yarn,

CC woven or non-woven fabric, for the degradation of plant material e.g.  
CC recycled waste paper, mechanical paper-making pulps or fibres subjected  
CC to retting process, for preparing animal feed and for processing wine or  
CC juice. DNA encoding the enzyme can also be used for the production of  
CC transgenic plants  
XX  
XX Sequence 221 AA.

Query Match	14.8%;	Score 341;	DB 2;	Length 221;
Best Local Similarity	41.7%;	Pred. No. 1.4e-15;		
Matches	73;	Conservative 33;	Mismatches 59;	Indels 10; Gaps 2
QY	254	VLHDTITVKAGVDFGKGTQTACSEIGDGGQSENQKPLFI	BDGASLKNVTGDDGADG	313
Db	31	VVKHIVVEKGTQYDGKRUIIAGPEIGDGSQREDQKPI	FVBDGATLKNVVLGPAADG	90
QY	314	IHLVGDAKIDNLHVNTVGEDAITVYKPSAGKSHVEITN	SSPEHSDKILQLNADNLNV	373
Db	91	VHTYGNASINNVMVEDGADALTVK----	SEGSVTINGSARLADKIFQINKASTFTV	145
QY	374	DNVKAADPGTFTVNTGGQGNWDNLNLSHISAEDGKPS	FVKDSEGLNVTNTDISL	428
Db	146	KNFTADQGGKFTIRLGSTFFKAVVINDCNTITNMKEAIF	RTDS-----STGSVTM	195

RESULT 18	
ADK48880	
ID	ADK48880 standard; protein; 221 AA.
XX	
XX	
AC	ADK48880;
XX	
XX	
DT	06-MAY-2004 (first entry)
XX	
XX	
DE	Bacillus pectic acid lyase protein SeqID 1.
XX	
XX	
KW	pectic acid lyase; enzyme; A-type proto-pectinase; washing agent;
KW	detergent; fibre processing agent; food-processing agent.
XX	
XX	
OS	Bacillus sp.
XX	
XX	
PN	JP2003250571-A.
XX	
XX	
PD	09-SEP-2003.

Claim 1; SEQ ID NO 1; 9pp; Japanese.

Query Match 14.8%; Score 341; DB 8; Length 221;  
Best Local Similarity 41.7%; Pred. No. 1.4e-15;

Matches	73;	Conservative	33;	Mismatches	59;	Indels	10;	Gaps	2;
QY	254	VLHDTTIVKAGOVFGKQGTFTAGSELGDCGQSENQKPLFILEDGASLKNVTWGGDGDG	313						
Db	31	VHKTIIVKGGQTYDGKGRLIAGPELGDGSGQREDQKPIFKVEDGATLKNVILGAPAADG	90						
QY	314	IHLYGDAKIDNLHVTVNGVEDAITVKPNSAGKKSHVETITNSFEHSDKILQLNADTNLSV	373						
Db	91	VHTYGNASINNVWEDVGEDALTVK-----SEGSVTINGGSARLAADKIFQINKASTFTV	145						
QY	374	DNVKAADGTFTVRNVGGQGGWMDNLSHISAEGCKFKFSKSDSEGLNVNTSDISL	428						
Db	146	KNFTADQGGKFIROLGSGSTFKAVVNIQNTITNNKKAIFRTDS-----STSSVTM	195						
RESULT 19									
ABU36481									
ID	ABU36481	standard; protein; 1306 AA.							
XX	AC	ABU36481;							
XX	AC	ABU36481;							
DT	19-JUN-2003	(first entry)							
XX	XX								
DE	Protein encoded by Prokaryotic essential gene #22008.								
XX	XX	Antisense; prokaryotic essential gene; cell proliferation; drug design.							
KW	XX	Mycobacterium tuberculosis.							
OS	XX	WO200277183-A2.							
FN	XX	03-OCT-2002.							
PD	XX	21-MAR-2002; 2002WO-US0009107.							
PF	XX	21-MAR-2001; 2001US-00815242.							
PR	XX	06-SEP-2001; 2001US-00948993.							
PR	XX	25-OCT-2001; 2001US-0342923P.							
PR	XX	08-FEB-2002; 2002US-00072851.							
PR	XX	06-MAR-2002; 2002US-0362699P.							
PA	XX	(ELIT-) ELITRA PHARM INC.							
PI	XX	Wang L, Zamudio C, Malone C, Haselbeck R, Ohlseen XL, Zyskind JW;							
PI	XX	Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;							
XX	XX	WFI; 2003-029926/02.							
DR	XX	N-PSDB; ACA40351.							
XX	XX	New antisense nucleic acids, useful for identifying proteins or screening							
PT	XX	for homologous nucleic acids required for cellular proliferation to							
PT	XX	isolate candidate molecules for rational drug discovery programs.							
XX	XX	Claim 25; SEQ ID NO 64405; 1766pp; English.							



CC product is overexpressed or underexpressed; (12) determining the extent  
 CC to which each of the strains is present in a culture or collection of  
 CC strains; or (13) identifying the target of a compound that inhibits the  
 CC proliferation of an organism. The antisense nucleic acids are useful for  
 CC identifying proteins or screening for homologous nucleic acids required  
 CC for cellular proliferation to isolate candidate molecules for rational  
 CC drug discovery programs, or for screening homologous nucleic acids  
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of  
 CC the target prokaryotic essential genes. Note: The sequence data for this  
 CC patent did not form part of the printed specification, but was obtained  
 CC in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 1306 AA;

Query Match 8.6%; Score 199.5; DB 6; Length 1306;  
 Best Local Similarity 26.0%; Pred. No. 8e-05;  
 Matches 88; Conservative 16; Mismatches 122; Indels 113; Gaps 11;

QY 8 NNTSSPGLFQSGDNLGHNANSALGQQPDRQTIEQMAQLLAKLLSLSPQSGNAA 67  
 DB 754 NGVAGSQPGAGGCGGTGGVCGGCGIDGADGAT----- 789  
 QY 68 TGAGGNDQTTGCVAGLNGKCTAGTTQSDSQNMLSEMNGNLDQAITPDQGGGQIG 127  
 DB 790 --AGARGQDGGAGGAGGCGRGCTGTP-----GGAGPAGTGTGQAGGNG 832  
 QY 128 DNPLLKAMLLIARMWGDGSDQFGPCTGNNSASGTSSTSGGSPFNDLSGKA-PSGNSP 186  
 DB 833 -----GSGGTGDPDGGNGANGSVFTNNGTGGNGGNGAGPAGG 875  
 QY 187 SGNYSVPSTFSPSTPTSP-----DFPSSPTKAAG-----G 221  
 DB 876 SGAG--STFG---ATGSSSIHVNGGNGGNGGNGGNGGNGGNGGNGGNGG 929  
 QY 222 STPTVTHPPVDSAGTACAGNSAFTAGANQTVLHDTITVKAQVFDGKGQFTTAGSELG 281  
 DB 930 SGAGGHHGNGGNASRMGMDGDTGTGAGGN-----AGQING-----GAGGNGG 973  
 QY 282 DGQSQENQPLFL-----EDGSLKNVTMGDDGADG 313  
 DB 974 DGTGSDGNPGLTSGGRRGGDGGVGGGGSVAGDGDG 1012

RESULT 20  
 ABU36802  
 ID ABU36802 standard; protein; 694 AA.  
 AC ABU36802;  
 DT 19-JUN-2003 (first entry)  
 XX  
 DE Protein encoded by Prokaryotic essential gene #22329.  
 XX  
 KW Antisense; prokaryotic essential gene; cell proliferation; drug design.  
 XX  
 OS Mycobacterium tuberculosis.  
 XX  
 PN WO20027183-A2.  
 XX  
 PD 03-OCT-2002.  
 XX  
 PF 21-MAR-2002; 2002WO-US009107.  
 XX  
 PR 21-MAR-2001; 2001US-00815242.  
 PR 06-SEP-2001; 2001US-00948993.  
 PR 25-OCT-2001; 2001US-0342923P.  
 PR 08-FEB-2002; 2002US-00072851.  
 PR 06-MAR-2002; 2002US-0362699P.  
 XX  
 PA (ELIT-) ELITRA PHARM INC.  
 XX

PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;  
 PI Wall D, Trawick JD, Carr GU, Yamamoto R, Foreyth RA, Xu HH;  
 XX WPI; 2003-029926/02.  
 DR N-PSDB; ACA40672.  
 XX  
 PT New antisense nucleic acids, useful for identifying proteins or screening  
 PT for homologous nucleic acids required for cellular proliferation to  
 PT isolate candidate molecules for rational drug discovery programs.  
 XX  
 XX Claim 25; SEQ ID NO 64726; 1766pp; English.  
 PS  
 CC The invention relates to an isolated nucleic acid comprising any one of  
 CC the 6213 antisense sequences given in the specification where expression  
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
 CC (1) a vector comprising a promoter operably linked to the nucleic acid  
 CC encoding a polypeptide whose expression is inhibited by the antisense  
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
 CC polypeptide or its fragment whose expression is inhibited by the  
 CC antisense nucleic acid; (4) an antibody capable of specifically binding  
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
 CC proliferation or the activity of a gene in an operon required for  
 CC proliferation; (7) identifying a compound that influences the activity of  
 CC the gene product or that has an activity against a biological pathway  
 CC required for proliferation, or that inhibits cellular proliferation; (8)  
 CC identifying a gene required for cellular proliferation or the biological  
 CC pathway in which a proliferation-required gene or its gene product lies  
 CC or a gene on which the test compound that inhibits proliferation of an  
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
 CC compound's activity; (11) a culture comprising strains in which the gene  
 CC product is overexpressed or underexpressed; (12) determining the extent  
 CC to which each of the strains is present in a culture or collection of  
 CC strains; or (13) identifying the target of a compound that inhibits the  
 CC proliferation of an organism. The antisense nucleic acids are useful for  
 CC identifying proteins or screening for homologous nucleic acids required  
 CC for cellular proliferation to isolate candidate molecules for rational  
 CC drug discovery programs, or for screening homologous nucleic acids  
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of  
 CC the target prokaryotic essential genes. Note: The sequence data for this  
 CC patent did not form part of the printed specification, but was obtained  
 CC in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 694 AA;

Query Match 8.4%; Score 193; DB 6; Length 694;  
 Best Local Similarity 25.4%; Pred. No. 0.0001;  
 Matches 86; Conservative 28; Mismatches 140; Indels 84; Gaps 14;

QY 18 QSGDNLGAGH-NANSALGQQPDRQTIEQMAQLLAKLLSLSPQSGNNAATGAGNDQT 76  
 DB 256 QTGGDGTGGHGTAGTPTGTGCGDGT-----ATAGSKATGAGGDDGT 300  
 QY 77 TGVNAGLNGKGTAGTTQSDSQNMLSEMNGNLDQAITPDQGGGQIGDNPLLKAML 136  
 DB 301 AAAGGGGNGGCGGVA---QGDIAAFGCGDGGSDGVAAGSGGSGGAGGG---AFV 352  
 QY 137 KLIARMWGDGSDQFGPCTGNNSASGTSSTSGGSPFNDLSGKK--APSGNSPSGNYSVP 193  
 DB 353 HIATATSTGSGGFG--GNGAASASGAGGAGGAGGAGGAGGLLFDGCGNGGAGGAGGI 410  
 QY 194 STFPSPTPTSPSTPLDFFSPSTKAAGSTPVT--DHDPD-----VGSAG 236  
 DB 411 GG-----DGATGGGSGGNGAGIARFSDPPEAPDPVVVGKGGGKGGSG 456  
 QY 237 IGAGNSVAFSTAGANQTVLHDTITVKAQVFDGKGQFTTAGSELGCGGSENGKPLFILE 296  
 DB 457 LGVGCAGGTGCGAGNG-----GAGGLLFGNGN---GGNAGAGG-----D 493  
 QY 297 DGASLKNVTMGDDGADGIHLYGDAKIDNL-HVTNVCED 333  
 DB 494 GGAGVAGVGNGGCGGCTATFHDPPVAGVWAVGGVGD 531

Thu Jan 27 12:40:01 2005

us-09-879-248-6.rag

Page 14

Search completed: January 25, 2005, 12:06:04  
Job time : 122 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2005, 12:04:41 ; Search time 58 Seconds

(without alignments)  
2784.420 Million cell updates/sec

Title: US-09-879-248-6

Perfect score: 2310

Sequence: 1 MSILLNNTSSPGLFQSG.....LGDVENVKVPMSANKVAE 447

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1608061 seqs, 361289386 residues

Total number of hits satisfying chosen parameters: 1608061

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.\*
- 19: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 20: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2310	100.0	447	9	US-09-835-684-5
2	2310	100.0	447	9	US-09-880-371-5
3	2310	100.0	447	9	US-09-879-248-6
4	2310	100.0	447	14	US-10-010-390-5
5	2310	100.0	447	15	US-10-441-736-6
6	2310	100.0	447	17	US-10-847-142-5
7	559	24.2	424	9	US-09-835-684-9
8	559	24.2	424	9	US-09-880-371-9
9	559	24.2	424	9	US-09-879-248-14
10	559	24.2	424	14	US-10-010-390-9
11	559	24.2	424	15	US-10-441-736-14
12	559	24.2	424	17	US-10-847-142-9
13	256	11.1	276	14	US-10-156-761-13910

14	199.5	8.6	1306	15	US-10-282-122A-64405	Sequence 64405, A
15	193	8.4	594	15	US-10-282-122A-64726	Sequence 64726, A
16	191.5	8.3	584	15	US-10-282-122A-50604	Sequence 50604, A
17	190.5	8.2	226	17	US-10-425-115-229387	Sequence 229387, A
18	188	8.1	639	15	US-10-282-122A-64609	Sequence 64609, A
19	186.5	8.1	591	15	US-10-282-122A-64363	Sequence 64363, A
20	185.5	8.0	1381	15	US-10-282-122A-64895	Sequence 64895, A
21	185	8.0	355	17	US-10-425-115-313677	Sequence 313677, A
22	183	7.9	588	15	US-10-282-122A-64869	Sequence 64869, A
23	182	7.9	606	15	US-10-282-122A-64464	Sequence 64464, A
24	181.5	7.9	484	15	US-09-820-843A-19	Sequence 19, Appl
25	181.5	7.9	484	15	US-10-282-122A-64867	Sequence 64867, A
26	180.5	7.8	1079	10	US-09-820-843A-20	Sequence 20, Appl
27	176.5	7.6	562	15	US-10-282-122A-64514	Sequence 64514, A
28	174	7.5	615	15	US-10-282-122A-64786	Sequence 64786, A
29	174	7.5	2042	14	US-10-192-584-6	Sequence 6, Appl
30	172.5	7.5	584	15	US-10-282-122A-64903	Sequence 64903, A
31	172	7.4	3507	14	US-10-369-493-5784	Sequence 5784, Ap
32	171.5	7.4	667	15	US-10-282-122A-64494	Sequence 64494, A
33	171	7.4	505	15	US-10-282-122A-62341	Sequence 62341, A
34	169.5	7.3	514	15	US-10-282-122A-64606	Sequence 64606, A
35	169	7.3	525	15	US-10-282-122A-64763	Sequence 64763, A
36	167.5	7.3	1011	15	US-10-282-122A-64589	Sequence 64589, A
37	166.5	7.2	518	15	US-10-282-122A-50634	Sequence 50634, A
38	166.5	7.2	532	15	US-10-282-122A-64658	Sequence 64658, A
39	166	7.2	461	15	US-10-282-122A-64750	Sequence 64750, A
40	166	7.2	603	15	US-10-282-122A-64537	Sequence 64537, A
41	166	7.2	767	15	US-10-282-122A-48384	Sequence 48384, A
42	165.5	7.2	344	9	US-09-086-118-27	Sequence 27, Appl
43	165.5	7.2	344	9	US-09-835-684-11	Sequence 11, Appl
44	165.5	7.2	344	9	US-09-880-371-11	Sequence 11, Appl
45	165.5	7.2	344	9	US-09-879-248-15	Sequence 15, Appl
46	165.5	7.2	344	9	US-09-770-693-7	Sequence 7, Appl
47	165.5	7.2	344	9	US-09-766-348-7	Sequence 7, Appl
48	165.5	7.2	344	14	US-10-034-158-7	Sequence 11, Appl
49	165.5	7.2	344	14	US-10-010-390-11	Sequence 11, Appl
50	165.5	7.2	344	14	US-10-387-806-27	Sequence 27, Appl
51	165.5	7.2	344	14	US-10-441-736-15	Sequence 15, Appl
52	165.5	7.2	344	17	US-10-847-142-11	Sequence 11, Appl
53	163	7.1	618	15	US-10-282-122A-64608	Sequence 64608, A
54	162.5	7.0	1017	14	US-10-369-493-20166	Sequence 20166, A
55	161.5	7.0	576	15	US-10-282-122A-64547	Sequence 64547, A
56	161.5	7.0	837	15	US-10-282-122A-64362	Sequence 64362, A
57	161.5	7.0	1215	14	US-10-369-493-10073	Sequence 10073, A
58	161	7.0	1601	15	US-10-263-929-132	Sequence 132, App
59	161	7.0	1601	15	US-10-267-502-241	Sequence 241, App
60	160.5	6.9	957	15	US-10-282-122A-64361	Sequence 64361, A
61	159.5	6.9	778	15	US-10-282-122A-64751	Sequence 64751, A
62	159	6.9	714	14	US-10-032-585-7539	Sequence 7539, Ap
63	159	6.9	3286	15	US-10-282-122A-49697	Sequence 49697, A
64	157	6.8	491	15	US-10-282-122A-64558	Sequence 64558, A
65	157	6.8	979	15	US-10-282-122A-47372	Sequence 47372, A
66	156	6.8	2039	14	US-10-192-584-7	Sequence 7, Appl
67	155.5	6.7	2353	9	US-09-797-862-33	Sequence 33, Appl
68	155	6.7	1872	15	US-10-282-122A-55254	Sequence 55254, A
69	154.5	6.7	354	15	US-09-820-843A-21	Sequence 21, Appl
70	154.5	6.7	630	14	US-10-128-714-8597	Sequence 8597, Ap
71	154	6.7	334	15	US-10-282-122A-47939	Sequence 47939, A
72	154	6.7	1099	15	US-10-282-122A-48220	Sequence 48220, A
73	153.5	6.6	597	9	US-09-793-306-146	Sequence 146, App
74	153.5	6.6	597	15	US-10-320-797-3261	Sequence 3261, Ap
75	153	6.6	611	14	US-10-086-510-4	Sequence 4, Appl
76	153	6.6	1466	15	US-10-402-089-4	Sequence 4, Appl
77	153	6.6	1466	15	US-10-402-089-6	Sequence 6, Appl
78	153	6.6	1466	15	US-10-403-072A-4	Sequence 4, Appl
79	153	6.6	1466	15	US-10-402-072A-6	Sequence 6, Appl
80	153	6.6	1770	9	US-09-841-132-444	Sequence 444, App
81	153	6.6	1770	16	US-10-467-534-21	Sequence 21, Appl
82	153	6.6	1770	17	US-10-872-155-444	Sequence 444, App
83	152.5	6.6	580	16	US-10-367-094-3	Sequence 3, Appl
84	152.5	6.6	821	14	US-10-292-798-1370	Sequence 1370, Ap
85	152.5	6.6	852	15	US-10-282-122A-43144	Sequence 43144, A
86	152.5	6.6	923	15	US-10-282-122A-64474	Sequence 64474, A

87 152 6.6 469 15 US-10-282-122A-45275, A Sequence 45275, A  
88 152 6.6 1082 14 US-10-369-493-17890 Sequence 17890, A  
89 152 6.6 3619 15 US-10-282-122A-67768 Sequence 67768, A  
90 151 6.5 256 10 US-09-820-843A-18 Sequence 18, Appl  
91 151 6.5 403 9 US-09-086-118-23 Sequence 23, Appl  
92 151 6.5 403 9 US-09-835-684-3 Sequence 3, Appl  
93 151 6.5 403 9 US-09-880-371-3 Sequence 3, Appl  
94 151 6.5 403 9 US-09-879-248-3 Sequence 3, Appl  
95 151 6.5 403 9 US-09-770-693-3 Sequence 3, Appl  
96 151 6.5 403 9 US-09-766-348-3 Sequence 3, Appl  
97 151 6.5 403 14 US-10-034-158-3 Sequence 3, Appl  
98 151 6.5 403 14 US-10-010-390-3 Sequence 3, Appl  
99 151 6.5 403 14 US-10-387-806-23 Sequence 23, Appl  
100 151 6.5 403 15 US-10-441-736-3 Sequence 3, Appl

## ALIGNMENTS

## RESULT 1

US-09-835-684-5  
; Sequence 5, Application US/09835684  
; Patent No. US20020019337A1  
; GENERAL INFORMATION:  
; APPLICANT: Wei, Zhong-Min  
; APPLICANT: Remick, Dean  
; TITLE OF INVENTION: TREATMENT OF FRUITS OR VEGETABLES WITH HYPERSENSITIVE  
; TITLE OF INVENTION: RESPONSE ELICITOR TO CONTROL POSTHARVEST DISEASE OR  
; TITLE OF INVENTION: DESICATION  
; FILE REFERENCE: 21829/71  
; CURRENT APPLICATION NUMBER: US/09/835,684  
; CURRENT FILING DATE: 2001-04-16  
; PRIOR APPLICATION NUMBER: 60/198,359  
; PRIOR FILING DATE: 2000-04-19  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 5  
; LENGTH: 447  
; TYPE: PRT  
; ORGANISM: Erwinia amylovora  
US-09-835-684-5

Query Match 100.0%; Score 2310; DB 9; Length 447;  
Best Local Similarity 100.0%; Pred. No. 1.4e-151;  
Matches 447; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MSILTLNNNTSSPGLFQSGDNGLGHNANSALGQOPIDRQTIEQMAQLLAELLKSLLS 60  
DB 1 MSILTLNNNTSSPGLFQSGDNGLGHNANSALGQOPIDRQTIEQMAQLLAELLKSLLS 60  
QY 61 POSGNAATGAGNDQTTGCVGNAGLNGRKTAGTTPOSQSONMLSEWNGNGLDQAITPDG 120  
DB 61 POSGNAATGAGNDQTTGCVGNAGLNGRKTAGTTPOSQSONMLSEWNGNGLDQAITPDG 120  
QY 121 QGGQIGDNPGLKAMKLIARMMDGQSDQFGQGTGNNSASSGTSSSGSGSPFNLDLGGKA 180  
DB 121 QGGQIGDNPGLKAMKLIARMMDGQSDQFGQGTGNNSASSGTSSSGSGSPFNLDLGGKA 180  
QY 181 PSNGSPSGNYSPTVSTFSPSTPTSPPLDPPSSPTKAAGGSTPVTDHPDPVGSAGIGAG 240  
DB 181 PSNGSPSGNYSPTVSTFSPSTPTSPPLDPPSSPTKAAGGSTPVTDHPDPVGSAGIGAG 240  
QY 241 NSVAFTSAGANQTVLHDTITVKAGQVDFGKGTFTAGSELGQSGSENQKPLFILEDGA 300  
DB 241 NSVAFTSAGANQTVLHDTITVKAGQVDFGKGTFTAGSELGQSGSENQKPLFILEDGA 300  
QY 301 LKNVTMGDDGADGHIHYGDAKIDNLHVTNVGDAITVKPNSAGKSHVEITNSSFEHASD 360  
DB 301 LKNVTMGDDGADGHIHYGDAKIDNLHVTNVGDAITVKPNSAGKSHVEITNSSFEHASD 360  
QY 361 KILQLNADTNLSVDNVKAKDFGTFTVTNGGQGNWDLNLSHISAEDGKFSFVKSDSEGLN 420  
DB 361 KILQLNADTNLSVDNVKAKDFGTFTVTNGGQGNWDLNLSHISAEDGKFSFVKSDSEGLN 420

## RESULT 3

US-09-879-248-6  
; Sequence 6, Application US/09879248  
; Patent No. US20020062500A1  
; GENERAL INFORMATION:  
; APPLICANT: Fan, Hao  
; APPLICANT: Wei, Zhong-Min

Db 361 KILQLNADTNLSVDNVKAKDFGTFTVTNGGQGNWDLNLSHISAEDGKFSFVKSDSEGLN 420  
QY 421 VNTSDISLGDVENHYKVPMSANLKVAE 447  
Db 421 VNTSDISLGDVENHYKVPMSANLKVAE 447  
RESULT 2  
US-09-880-371-5  
; Sequence 5, Application US/09880371  
; Patent No. US20020059658A1  
; GENERAL INFORMATION:  
; APPLICANT: Wei, Zhong-Min  
; APPLICANT: DeRoche, Jay  
; TITLE OF INVENTION: METHODS OF IMPROVING THE EFFECTIVENESS OF TRANSGENIC  
; TITLE OF INVENTION: PLANTS  
; FILE REFERENCE: 21829/91  
; CURRENT APPLICATION NUMBER: US/09/880,371  
; CURRENT FILING DATE: 2001-06-13  
; PRIOR APPLICATION NUMBER: 60/211,585  
; PRIOR FILING DATE: 2000-06-15  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 5  
; LENGTH: 447  
; TYPE: PRT  
; ORGANISM: Erwinia amylovora  
US-09-880-371-5

Query Match 100.0%; Score 2310; DB 9; Length 447;  
Best Local Similarity 100.0%; Pred. No. 1.4e-151;  
Matches 447; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MSILTLNNNTSSPGLFQSGDNGLGHNANSALGQOPIDRQTIEQMAQLLAELLKSLLS 60  
DB 1 MSILTLNNNTSSPGLFQSGDNGLGHNANSALGQOPIDRQTIEQMAQLLAELLKSLLS 60  
QY 61 POSGNAATGAGNDQTTGCVGNAGLNGRKTAGTTPOSQSONMLSEWNGNGLDQAITPDG 120  
DB 61 POSGNAATGAGNDQTTGCVGNAGLNGRKTAGTTPOSQSONMLSEWNGNGLDQAITPDG 120  
QY 121 QGGQIGDNPGLKAMKLIARMMDGQSDQFGQGTGNNSASSGTSSSGSGSPFNLDLGGKA 180  
DB 121 QGGQIGDNPGLKAMKLIARMMDGQSDQFGQGTGNNSASSGTSSSGSGSPFNLDLGGKA 180  
QY 181 PSNGSPSGNYSPTVSTFSPSTPTSPPLDPPSSPTKAAGGSTPVTDHPDPVGSAGIGAG 240  
DB 181 PSNGSPSGNYSPTVSTFSPSTPTSPPLDPPSSPTKAAGGSTPVTDHPDPVGSAGIGAG 240  
QY 241 NSVAFTSAGANQTVLHDTITVKAGQVDFGKGTFTAGSELGQSGSENQKPLFILEDGA 300  
DB 241 NSVAFTSAGANQTVLHDTITVKAGQVDFGKGTFTAGSELGQSGSENQKPLFILEDGA 300  
QY 301 LKNVTMGDDGADGHIHYGDAKIDNLHVTNVGDAITVKPNSAGKSHVEITNSSFEHASD 360  
DB 301 LKNVTMGDDGADGHIHYGDAKIDNLHVTNVGDAITVKPNSAGKSHVEITNSSFEHASD 360  
QY 361 KILQLNADTNLSVDNVKAKDFGTFTVTNGGQGNWDLNLSHISAEDGKFSFVKSDSEGLN 420  
DB 361 KILQLNADTNLSVDNVKAKDFGTFTVTNGGQGNWDLNLSHISAEDGKFSFVKSDSEGLN 420  
QY 421 VNTSDISLGDVENHYKVPMSANLKVAE 447  
DB 421 VNTSDISLGDVENHYKVPMSANLKVAE 447

; TITLE OF INVENTION: HYPERSENSITIVE RESPONSE ELICITING DOMAINS AND USE
; FILE REFERENCE: 21829/81
; CURRENT APPLICATION NUMBER: US/09/879, 248
; CURRENT FILING DATE: 2001-06-12
; PRIOR APPLICATION NUMBER: 60/212,211
; PRIOR FILING DATE: 2000-06-16
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 447
; TYPE: PRT
; ORGANISM: Erwinia amylovora
US-09-879-248-6
Query Match 100.0%; Score 2310; DB 9; Length 447;
Best Local Similarity 100.0%; Pred. No. 1.4e-151;
Matches 447; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSILTLNNNTSSPGLFQSGDNGLGHNANSALGQPIDRQTIQMAQLLAELLKSLLS 60
DB 1 MSILTLNNNTSSPGLFQSGDNGLGHNANSALGQPIDRQTIQMAQLLAELLKSLLS 60
QY 61 POSGNAATGAGNDQTTGVGNAGLNGRKGTAGTTTQSDSQNNMLSEMGNNGLDQAITPDG 120
DB 61 POSGNAATGAGNDQTTGVGNAGLNGRKGTAGTTTQSDSQNNMLSEMGNNGLDQAITPDG 120
QY 121 QGGGQIGDNPGLKAMUKLIARMMDGSDQFGQGTGNNSSASCTSSGGSPFNDLSGGKA 180
DB 121 QGGGQIGDNPGLKAMUKLIARMMDGSDQFGQGTGNNSSASCTSSGGSPFNDLSGGKA 180
QY 181 PSNPSGNGYSPVSTPSTPTSPPLDPSSPTKAAGSTPVTDHPDPVGSAGIGAG 240
DB 181 PSNPSGNGYSPVSTPSTPTSPPLDPSSPTKAAGSTPVTDHPDPVGSAGIGAG 240
QY 241 NSVAFTSAGANQTVLHDTITVKAGQVFDGKQGTFTAGSELGCGQSENOKPLFILEDGAS 300
DB 241 NSVAFTSAGANQTVLHDTITVKAGQVFDGKQGTFTAGSELGCGQSENOKPLFILEDGAS 300
QY 301 LKQVMTGDDGADGHIHYLGDAKIDNLHVTNVGDAITVKPNSACKKSHVEITNSSFEHASD 360
DB 301 LKQVMTGDDGADGHIHYLGDAKIDNLHVTNVGDAITVKPNSACKKSHVEITNSSFEHASD 360
QY 361 KILQLNADTNLSVDNVKAKDFGTFVRTNGQQGNWDLNLSHISAEDGKFSFVKSDSEGLN 420
DB 361 KILQLNADTNLSVDNVKAKDFGTFVRTNGQQGNWDLNLSHISAEDGKFSFVKSDSEGLN 420
QY 421 VNTSDISLGDVENHYKVPMSANLKVAE 447
DB 421 VNTSDISLGDVENHYKVPMSANLKVAE 447
RESULT 4
US-10-010-390-5
; Sequence 5, Application US/10010390
; Publication No. US20030104979A1
; GENERAL INFORMATION:
; APPLICANT: Wei, Zhong-Min
; APPLICANT: Leon, Ernesto
; APPLICANT: Oviedo, Agustín
; TITLE OF INVENTION: METHODS OF INHIBITING DESICCATION OF CUTTINGS REMOVED
; FILE REFERENCE: 21829/111
; CURRENT APPLICATION NUMBER: US/10/010,390
; PRIOR FILING DATE: 2001-11-05
; PRIOR FILING DATE: 2000-11-13
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 447
; TYPE: PRT
; ORGANISM: Erwinia amylovora

US-10-010-390-5
Query Match 100.0%; Score 2310; DB 14; Length 447;
Best Local Similarity 100.0%; Pred. No. 1.4e-151;
Matches 447; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSILTLNNNTSSPGLFQSGDNGLGHNANSALGQPIDRQTIQMAQLLAELLKSLLS 60
DB 1 MSILTLNNNTSSPGLFQSGDNGLGHNANSALGQPIDRQTIQMAQLLAELLKSLLS 60
QY 61 POSGNAATGAGNDQTTGVGNAGLNGRKGTAGTTTQSDSQNNMLSEMGNNGLDQAITPDG 120
DB 61 POSGNAATGAGNDQTTGVGNAGLNGRKGTAGTTTQSDSQNNMLSEMGNNGLDQAITPDG 120
QY 121 QGGGQIGDNPGLKAMUKLIARMMDGSDQFGQGTGNNSSASCTSSGGSPFNDLSGGKA 180
DB 121 QGGGQIGDNPGLKAMUKLIARMMDGSDQFGQGTGNNSSASCTSSGGSPFNDLSGGKA 180
QY 181 PSNPSGNGYSPVSTPSTPTSPPLDPSSPTKAAGSTPVTDHPDPVGSAGIGAG 240
DB 181 PSNPSGNGYSPVSTPSTPTSPPLDPSSPTKAAGSTPVTDHPDPVGSAGIGAG 240
QY 241 NSVAFTSAGANQTVLHDTITVKAGQVFDGKQGTFTAGSELGCGQSENOKPLFILEDGAS 300
DB 241 NSVAFTSAGANQTVLHDTITVKAGQVFDGKQGTFTAGSELGCGQSENOKPLFILEDGAS 300
QY 301 LKQVMTGDDGADGHIHYLGDAKIDNLHVTNVGDAITVKPNSACKKSHVEITNSSFEHASD 360
DB 301 LKQVMTGDDGADGHIHYLGDAKIDNLHVTNVGDAITVKPNSACKKSHVEITNSSFEHASD 360
QY 361 KILQLNADTNLSVDNVKAKDFGTFVRTNGQQGNWDLNLSHISAEDGKFSFVKSDSEGLN 420
DB 361 KILQLNADTNLSVDNVKAKDFGTFVRTNGQQGNWDLNLSHISAEDGKFSFVKSDSEGLN 420
QY 421 VNTSDISLGDVENHYKVPMSANLKVAE 447
DB 421 VNTSDISLGDVENHYKVPMSANLKVAE 447
RESULT 5
US-10-441-736-6
; Sequence 6, Application US/10441736
; Publication No. US20040016029A1
; GENERAL INFORMATION:
; APPLICANT: Wei, Zhong-Min
; APPLICANT: Schading, Richard L.
; TITLE OF INVENTION: HYPERSENSITIVE RESPONSE ELICITOR-INDUCED STRESS
; FILE REFERENCE: 21829/203 (BBC-003)
; CURRENT APPLICATION NUMBER: US/10/441,736
; CURRENT FILING DATE: 2003-05-20
; PRIOR APPLICATION NUMBER: 60/107,243
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: 09/431,614
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 447
; TYPE: PRT
; ORGANISM: Erwinia amylovora
US-10-441-736-6
Query Match 100.0%; Score 2310; DB 15; Length 447;
Best Local Similarity 100.0%; Pred. No. 1.4e-151;
Matches 447; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSILTLNNNTSSPGLFQSGDNGLGHNANSALGQPIDRQTIQMAQLLAELLKSLLS 60
DB 1 MSILTLNNNTSSPGLFQSGDNGLGHNANSALGQPIDRQTIQMAQLLAELLKSLLS 60
QY 61 POSGNAATGAGNDQTTGVGNAGLNGRKGTAGTTTQSDSQNNMLSEMGNNGLDQAITPDG 120

Db 61 POSNAATGAGGNDQTTGCVNAGGLNGRKGTTAGTTPQSDSQNMLSEMGNNGLDQAITPDG 120  
Qy 121 QGGGQIGDNPLLKAMLKLIARMDGQSDQFGQPGTGNNSASGTSSSGGSPFNDLSGGKA 180  
Db 121 QGGGQIGDNPLLKAMLKLIARMDGQSDQFGQPGTGNNSASGTSSSGGSPFNDLSGGKA 180  
Qy 181 PSGNSPGSNYPVSTFPSPPTPTGTPPLDPPSPPTKAAAGGSTPVTDHPDPVGSAGIGAG 240  
Db 181 PSGNSPGSNYPVSTFPSPPTPTGTPPLDPPSPPTKAAAGGSTPVTDHPDPVGSAGIGAG 240  
Qy 241 NSVAFTSAGANQTVLHDTITVKAGOVDFGKGTTFVTRTNGGQGNWDLNLSHISAEDCKFSFKVSDSEGLN 420  
Db 241 NSVAFTSAGANQTVLHDTITVKAGOVDFGKGTTFVTRTNGGQGNWDLNLSHISAEDCKFSFKVSDSEGLN 420  
Qy 301 LKNTVMGDDGADGIHLXGDAKIDNLHVTNVGDEDAITVKPNSAGKSHVEITNSSFHSHD 360  
Db 301 LKNTVMGDDGADGIHLXGDAKIDNLHVTNVGDEDAITVKPNSAGKSHVEITNSSFHSHD 360  
Qy 361 KILQNLADTNLSVDNVKAKDFGTFTVTRTNGGQGNWDLNLSHISAEDCKFSFKVSDSEGLN 420  
Db 361 KILQNLADTNLSVDNVKAKDFGTFTVTRTNGGQGNWDLNLSHISAEDCKFSFKVSDSEGLN 420  
Qy 421 VNTSDISLGDVENHYKVPMSANLKVAE 447  
Db 421 VNTSDISLGDVENHYKVPMSANLKVAE 447

RESULT 6  
US-10-847-142-5  
; Sequence 5, Application US/10847142  
; Publication No. US20040265442A1  
; GENERAL INFORMATION:  
; APPLICANT: Wei, Zhong-Min  
; APPLICANT: Qiu, Dewen  
; APPLICANT: Remick, Dean  
; TITLE OF INVENTION: TREATMENT OF FRUITS OR VEGETABLES WITH HYPERSENSITIVE  
; TITLE OF INVENTION: RESPONSE ELICITOR TO CONTROL POSTHARVEST DISEASE OR  
; TITLE OF INVENTION: DESICCATION  
; FILE REFERENCE: 21829/197  
; CURRENT APPLICATION NUMBER: US/10/847,142  
; CURRENT FILING DATE: 2004-05-17  
; PRIOR APPLICATION NUMBER: 60/198,359  
; PRIOR FILING DATE: 2000-04-19  
; PRIOR APPLICATION NUMBER: 09/835,684  
; PRIOR FILING DATE: 2001-04-16  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 5  
; LENGTH: 447  
; TYPE: PRT  
; ORGANISM: Erwinia amylovora  
US-10-847-142-5

Query Match 100.0%; Score 2310; DB 17; Length 447;  
Best Local Similarity 100.0%; Pred. No. 1.4e-151;  
Matches 447; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSILTLNNNTSSPGLFQSGGDNGLGHNANSALGQOPIDRQTIEQMAQLLAELLKSLLS 60  
Db 1 MSILTLNNNTSSPGLFQSGGDNGLGHNANSALGQOPIDRQTIEQMAQLLAELLKSLLS 60  
Qy 61 POSGNAATGAGGNDQTTGCVNAGGLNGRKGTTAGTTPQSDSQNMLSEMGNNGLDQAITPDG 120  
Db 61 POSGNAATGAGGNDQTTGCVNAGGLNGRKGTTAGTTPQSDSQNMLSEMGNNGLDQAITPDG 120  
Qy 121 QGGGQIGDNPLLKAMLKLIARMDGQSDQFGQPGTGNNSASGTSSSGGSPFNDLSGGKA 180  
Db 121 QGGGQIGDNPLLKAMLKLIARMDGQSDQFGQPGTGNNSASGTSSSGGSPFNDLSGGKA 180  
Qy 181 PSGNSPGSNYPVSTFPSPPTPTGTPPLDPPSPPTKAAAGGSTPVTDHPDPVGSAGIGAG 240  
Db 181 PSGNSPGSNYPVSTFPSPPTPTGTPPLDPPSPPTKAAAGGSTPVTDHPDPVGSAGIGAG 240

## RESULT 7

US-09-835-684-9  
; Sequence 9, Application US/09835684  
; Patent No. US20020019337A1  
; GENERAL INFORMATION:  
; APPLICANT: Wei, Zhong-Min  
; APPLICANT: Qiu, Dewen  
; APPLICANT: Remick, Dean  
; TITLE OF INVENTION: TREATMENT OF FRUITS OR VEGETABLES WITH HYPERSENSITIVE  
; TITLE OF INVENTION: RESPONSE ELICITOR TO CONTROL POSTHARVEST DISEASE OR  
; TITLE OF INVENTION: DESICCATION  
; FILE REFERENCE: 21829/71  
; CURRENT APPLICATION NUMBER: US/09/835,684  
; CURRENT FILING DATE: 2001-04-16  
; PRIOR APPLICATION NUMBER: 60/198,359  
; PRIOR FILING DATE: 2000-04-19  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 9  
; LENGTH: 424  
; TYPE: PRT  
; ORGANISM: Pseudomonas syringae  
US-09-835-684-9

Query Match 24.2%; Score 559; DB 9; Length 424;  
Best Local Similarity 36.1%; Pred. No. 1.9e-30;  
Matches 147; Conservative 54; Mismatches 132; Indels 74; Gaps 13;

Qy 37 QPIDRQTIEQMAQLLAELLKSL---LSPQSGNAATGAGGNDQTTGCVNAGGLNGRKGTTAG 93  
Db 72 KENDSQS--NIAKLISALIMSLLOMLTNSNKKQDTNQEPDSQAPFQNNGLG----- 122  
Qy 94 TTPQSDSQNMLSEMGNNGLDQAITPDGQGGQIGDNPLLKAMLKLIARMDGQSDQFGQP 153  
Db 123 -TPSADS-----GGGG-----TPDATGGGG-GDTP-----SATGGG 151  
Qy 154 GTGNNSASGTSSSGGSPFNDLSGGKAPSGNSPGSNYPVSTFPSPPTPTGTPPLDPPS 213  
Db 152 GDTPTATGCGGSGGGGTPPTATGGG---SGGTPTATGCGGSGGTPQITPQL-----A 200  
Qy 214 SPTKAAAGGSTPVTDHPDPVGSAGIAGNSVAPTSAGANQTVLHDTITVKAGOVDFGKGT 273  
Db 201 NPNRTSG-----TGVSVDTAGS---TQAGKINVVKDTIKVGADEVDFHGAT 245  
Qy 274 FTAGSELGDSQSNQKPLFILEDQASLKNTVMGDDGADGIHLXG---DAKIDNLHVTN 329  
Db 246 FTADKSMGNGDQGENQKPMFELAEATLKNVNLGENEVDGIHVAKAKNAQEVTTIDNVHAQ 305  
Qy 330 VGEDAITVKPNSAGKSHVEITNSSFHSHDQKIDNLADTNLSVDNVKAKDFGTFTVTRTNG 389  
Db 306 VGEDLITVKGGGGAATVNLNKNSSAKGADKKVQLNANTHLKIDNFKADDFGTMVTRNG 365  
Qy 390 GQQ--GNWDLNLSHISAEDCKFSFKVSDSEGLNNTSDISLGDVENHY 435  
Db 366 GKQFDDMSIELNGIEANHGKFAVKSDSDDLKLAATGNIAMTVDVKHAY 412

```

; LENGTH: 424
; TYPE: PRT
; ORGANISM: Pseudomonas syringae
US-09-879-248-14

Query Match 24.2%; Score 559; DB 9; Length 424;
Best Local Similarity 36.1%; Pred. No. 1.9e-30;
Matches 147; Conservative 54; Mismatches 132; Indels 74; Gaps 13;

QY 37 QPIDRTIEQMAQLLAELKSL---LSPQSGNAATGAGNDQTTGVGNAGLNGRKGATAG 93
Db 72 KPNDQS--NIAKLISALIMSLQMLTNSKKQDTNQEOPDSQAPFQNGGLG-----122
QY 94 TTPQSDSQNMLSEMGNGLDQAITPDGCGGQIGDNPPLKAMKLKLIARMMDGSDQFGQP 153
Db 123 -TPSADS-----GGGG-----TPDATGGG-GDTP-----SATGGG 151
QY 154 GTGNSSASCTSSGGSPFNDLGGKAPSGNSGVSPVSTFSPSTPTSPSPLDPPS 213
Db 152 GGDTPATGGGGGGGTPTATGG---SGTPTATGGGGGVTPTTQITPOL-----A 200
QY 214 SPTKAAAGSTPVTDPVGSAGIGAGNSVAFTSAGANQTVLHDTITVKAGQVDFGKGOT 273
Db 201 NPRTSG-----TGSVSDTAGS---TEQAGKINVVKDTIKVGAGEVDFDGHGAT 245
QY 274 FTAGSELGCGQSENQKPLFILEDGASLKNVTWGGDAGDIHLG-----DAKIDNLHVTN 329
Db 246 FTADKSMGNGDQENQKPMFELAEAGATLKNVNLGENEVDGIHVKAQNAQEVTTIDNVHAQN 305
QY 330 VGEDAITVKPNSAGKSHVEITNSSPEHASDKILOLNADTNLSVDNVKAKDFGTFVVRTNG 389
Db 306 VGEDLITVKGEGGAATVNLNKNSSAKGADDKVVQLNANTHLKIDNFKADDFGTFVVRTNG 365
QY 390 GQO-GNWDNLNLSHISAEDGKFSFKVSDSEGLNVNTSDISLGDVENHY 435
Db 366 GKQFDDMSIELNGIEANHGKALFKVSDSDDLKATGNIAMTDVKHAY 412

RESULT 8
US-09-880-371-9
; Sequence 9, Application US/09880371
; Patent No. US20020059658A1
; GENERAL INFORMATION:
; APPLICANT: Wei, Zhong-Min
; APPLICANT: Derocher, Jay
; TITLE OF INVENTION: METHODS OF IMPROVING THE EFFECTIVENESS OF TRANSGENIC
; FILE REFERENCE: 21829/91
; CURRENT APPLICATION NUMBER: US/09/880,371
; PRIOR FILING DATE: 2001-06-13
; PRIOR FILING DATE: 2000-06-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 424
; TYPE: PRT
; ORGANISM: Pseudomonas syringae
US-09-880-371-9

Query Match 24.2%; Score 559; DB 9; Length 424;
Best Local Similarity 36.1%; Pred. No. 1.9e-30;
Matches 147; Conservative 54; Mismatches 132; Indels 74; Gaps 13;

QY 37 QPIDRTIEQMAQLLAELKSL---LSPQSGNAATGAGNDQTTGVGNAGLNGRKGATAG 93
Db 72 KPNDQS--NIAKLISALIMSLQMLTNSKKQDTNQEOPDSQAPFQNGGLG-----122
QY 94 TTPQSDSQNMLSEMGNGLDQAITPDGCGGQIGDNPPLKAMKLKLIARMMDGSDQFGQP 153
Db 123 -TPSADS-----GGGG-----TPDATGGG-GDTP-----SATGGG 151
QY 154 GTGNSSASCTSSGGSPFNDLGGKAPSGNSGVSPVSTFSPSTPTSPSPLDPPS 213
Db 152 GGDTPATGGGGGGGTPTATGG---SGTPTATGGGGGVTPTTQITPOL-----A 200
QY 214 SPTKAAAGSTPVTDPVGSAGIGAGNSVAFTSAGANQTVLHDTITVKAGQVDFGKGOT 273
Db 201 NPRTSG-----TGSVSDTAGS---TEQAGKINVVKDTIKVGAGEVDFDGHGAT 245
QY 274 FTAGSELGCGQSENQKPLFILEDGASLKNVTWGGDAGDIHLG-----DAKIDNLHVTN 329
Db 246 FTADKSMGNGDQENQKPMFELAEAGATLKNVNLGENEVDGIHVKAQNAQEVTTIDNVHAQN 305
QY 330 VGEDAITVKPNSAGKSHVEITNSSPEHASDKILOLNADTNLSVDNVKAKDFGTFVVRTNG 389
Db 306 VGEDLITVKGEGGAATVNLNKNSSAKGADDKVVQLNANTHLKIDNFKADDFGTFVVRTNG 365
QY 390 GQO-GNWDNLNLSHISAEDGKFSFKVSDSEGLNVNTSDISLGDVENHY 435
Db 366 GKQFDDMSIELNGIEANHGKALFKVSDSDDLKATGNIAMTDVKHAY 412

RESULT 9
US-09-879-248-14
; Sequence 14, Application US/09879248
; Patent No. US20020062500A1
; GENERAL INFORMATION:
; APPLICANT: Fan, Hao
; APPLICANT: Fan, Hao
; TITLE OF INVENTION: HYPERSENSITIVE RESPONSE ELICITING DOMAINS AND USE
; FILE REFERENCE: 21829/81
; CURRENT APPLICATION NUMBER: US/09/879,248
; PRIOR FILING DATE: 2001-06-12
; PRIOR FILING DATE: 2000-06-16
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
```

```

; LENGTH: 424
; TYPE: PRT
; ORGANISM: Pseudomonas syringae
US-09-879-248-14

Query Match 24.2%; Score 559; DB 9; Length 424;
Best Local Similarity 36.1%; Pred. No. 1.9e-30;
Matches 147; Conservative 54; Mismatches 132; Indels 74; Gaps 13;

QY 37 QPIDRTIEQMAQLLAELKSL---LSPQSGNAATGAGNDQTTGVGNAGLNGRKGATAG 93
Db 72 KPNDQS--NIAKLISALIMSLQMLTNSKKQDTNQEOPDSQAPFQNGGLG-----122
QY 94 TTPQSDSQNMLSEMGNGLDQAITPDGCGGQIGDNPPLKAMKLKLIARMMDGSDQFGQP 153
Db 123 -TPSADS-----GGGG-----TPDATGGG-GDTP-----SATGGG 151
QY 154 GTGNSSASCTSSGGSPFNDLGGKAPSGNSGVSPVSTFSPSTPTSPSPLDPPS 213
Db 152 GGDTPATGGGGGGGTPTATGG---SGTPTATGGGGGVTPTTQITPOL-----A 200
QY 214 SPTKAAAGSTPVTDPVGSAGIGAGNSVAFTSAGANQTVLHDTITVKAGQVDFGKGOT 273
Db 201 NPRTSG-----TGSVSDTAGS---TEQAGKINVVKDTIKVGAGEVDFDGHGAT 245
QY 274 FTAGSELGCGQSENQKPLFILEDGASLKNVTWGGDAGDIHLG-----DAKIDNLHVTN 329
Db 246 FTADKSMGNGDQENQKPMFELAEAGATLKNVNLGENEVDGIHVKAQNAQEVTTIDNVHAQN 305
QY 330 VGEDAITVKPNSAGKSHVEITNSSPEHASDKILOLNADTNLSVDNVKAKDFGTFVVRTNG 389
Db 306 VGEDLITVKGEGGAATVNLNKNSSAKGADDKVVQLNANTHLKIDNFKADDFGTFVVRTNG 365
QY 390 GQO-GNWDNLNLSHISAEDGKFSFKVSDSEGLNVNTSDISLGDVENHY 435
Db 366 GKQFDDMSIELNGIEANHGKALFKVSDSDDLKATGNIAMTDVKHAY 412

RESULT 10
US-10-010-390-9
; Sequence 9, Application US/10010390
; Publication No. US20030104979A1
; GENERAL INFORMATION:
; APPLICANT: Wei, Zhong-Min
; APPLICANT: Leon, Ernesto
; APPLICANT: Oviedo, Agustín
; TITLE OF INVENTION: METHODS OF INHIBITING DESICCATION OF CUTTINGS REMOVED
; FILE REFERENCE: 21829/111
; CURRENT APPLICATION NUMBER: US/10/010,390
; PRIOR FILING DATE: 2001-11-05
; PRIOR FILING DATE: 2000-11-13
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 424
; TYPE: PRT
; ORGANISM: Pseudomonas syringae
US-10-010-390-9

Query Match 24.2%; Score 559; DB 14; Length 424;
Best Local Similarity 36.1%; Pred. No. 1.9e-30;
Matches 147; Conservative 54; Mismatches 132; Indels 74; Gaps 13;

QY 37 QPIDRTIEQMAQLLAELKSL---LSPQSGNAATGAGNDQTTGVGNAGLNGRKGATAG 93
Db 72 KPNDQS--NIAKLISALIMSLQMLTNSKKQDTNQEOPDSQAPFQNGGLG-----122
QY 94 TTPQSDSQNMLSEMGNGLDQAITPDGCGGQIGDNPPLKAMKLKLIARMMDGSDQFGQP 153
Db 123 -TPSADS-----GGGG-----TPDATGGG-GDTP-----SATGGG 151
```

```

QY 154 GTGNSSASGTSSTSSGGSPFNDLSGGKAPSGNSGNSYSPVSTFSPSTPTSPPLDPPS 213
Db 152 GGDTPATATGGGGGGGGTATATGGG---SGGTPATATGGGGGGTTPQITPQL-----A 200
QY 214 SPTKAAGGSTPTVTHDPDPVGSAGIGAGNSVAFTSAGANQTVLHDTITVKAGQVDFGKGT 273
Db 201 NPNRTSG-----TGSVSDTAGS-----TEQAGKINVVKDTIKVGAGEVDFGHGAT 245
QY 274 FTAGSELGGGQSGENKQPLFILEDGASIKNVMTGDDGADGHIHYG-----DAKIDNLHVTN 329
Db 246 FTADKSMGNGDQGENKQKPFELAEAGATLKNVNLGENEVDGIHVAKAKNAQEVITIDNVHAQN 305
QY 330 VGEDAITVKPNSAGKSHVEITNSSFEHASDKILQNLADTNLSVDNVKAKDFGTFVVRTNG 389
Db 306 VGEDLITVKEGGAAVTNLTNKNSSAKGADDKVQNLNANTHLKIDNFKADDFGTFVVRTNG 365
QY 390 GQO-GNWDNLNLSHISAEDGKFSFKVSDSEGLNVNTSDISLGDVENHY 435
Db 366 GKQFDDMSIELNGIEANHGKFAVSKSDSDDLKLTATGNIAMTVDVKHAY 412

RESULT 11
US-10-441-736-14
; Sequence 14, Application US/10441736
; Publication No. US20040016029A1
; GENERAL INFORMATION:
; APPLICANT: Wei, Zhong-Min
; APPLICANT: Schading, Richard L.
; TITLE OF INVENTION: HYPERSENSITIVE RESPONSE ELICITOR-INDUCED STRESS
; FILE REFERENCE: 21829/203 (EBC-003)
; CURRENT APPLICATION NUMBER: US/10/441,736
; PRIOR FILING DATE: 2003-05-20
; PRIOR APPLICATION NUMBER: 60/107,243
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: 09/431,614
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 424
; TYPE: PRT
; ORGANISM: Pseudomonas syringae
US-10-441-736-14

Query Match 24.2%; Score 559; DB 15; Length 424;
Best Local Similarity 36.1%; Pred. No. 1.9e-30;
Matches 147; Conservative 54; Mismatches 132; Indels 74; Gaps 13;

QY 37 QPIDRTQIEOMAQLLAELLKSL---LSPQSGNAATGAGGNDQTTGTVGNAGGLNKRKGTAG 93
Db 72 KPNDQS--NIAKLISALIMSLQMLTNSNKKQDTNQEQPDSQAPFQNNGLG-----122
QY 94 TTPQSDSQNMLSEMGNGLDQAITPDGQGGQIGDNPFLKAMLKLIARMMDGQSDQFGQP 153
Db 123 -TPSADS-----GGGG-----TPDATGGG-GDTP-----SATGG 151
QY 154 GTGNSSASGTSSTSSGGSPFNDLSGGKAPSGNSGNSYSPVSTFSPSTPTSPPLDPPS 213
Db 152 GGDTPATATGGGGGGGGTATATGGG---SGGTPATATGGGGGGTTPQITPQL-----A 200
QY 214 SPTKAAGGSTPTVTHDPDPVGSAGIGAGNSVAFTSAGANQTVLHDTITVKAGQVDFGKGT 273
Db 201 NPNRTSG-----TGSVSDTAGS-----TEQAGKINVVKDTIKVGAGEVDFGHGAT 245
QY 274 FTAGSELGGGQSGENKQPLFILEDGASIKNVMTGDDGADGHIHYG-----DAKIDNLHVTN 329
Db 246 FTADKSMGNGDQGENKQKPFELAEAGATLKNVNLGENEVDGIHVAKAKNAQEVITIDNVHAQN 305
QY 330 VGEDAITVKPNSAGKSHVEITNSSFEHASDKILQNLADTNLSVDNVKAKDFGTFVVRTNG 389
Db 306 VGEDLITVKEGGAAVTNLTNKNSSAKGADDKVQNLNANTHLKIDNFKADDFGTFVVRTNG 365

QY 390 GQO-GNWDNLNLSHISAEDGKFSFKVSDSEGLNVNTSDISLGDVENHY 435
Db 366 GKQFDDMSIELNGIEANHGKFAVSKSDSDDLKLTATGNIAMTVDVKHAY 412

RESULT 13
US-10-156-761-13910
; Sequence 13910, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI

```







GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 229387  
; LENGTH: 226  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)-(226)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_140794C.1.pep  
US-10-425-115-229387

Query Match 8.2%; Score 190.5; DB 17; Length 226;  
Best Local Similarity 26.6%; Pred. No. 2.8e-05;  
Matches 61; Conservative 41; Mismatches 82; Indels 45; Gaps 8;  
QY 199 PSTPTSPSPL---DFPSSPTKAAGGST-----PVTDPDPVGSAGIGAGNSVAF 245  
DB 13 PALSTAIVIPRQNGDPOGSTDSTQDGLGARPPRPFPVTKFPV-----58  
QY 246 TSAGANTVLHDTITVKAGQVFGK-----GOTTAGSELGDSQSENQKPLFILEDGL 301  
DB 59 ----TETVLKEMIAAAGQVFDGQMKRYGRVESCSEQAEGTDAD----AVFNLLPGATI 111  
QY 302 KNTWGGDAGDGHLYGDAKIDNLHVTNVCEDAITVKPSAGKSHVEITNSSFESASDK 361  
DB 112 RNVIIGKHQSEGHALGDWVENVWEDVCEADLT-----SKGLNTQLRVITGGARNATDK 167  
QY 362 ILQLNA-DTNLSVDNVKAKDFGTFVRTNGG--OQGNWDLNLSHISAEDG 407  
DB 168 IFQNSLGGKWNITGTFVAENFGTFYXSCGNXHQAKENVITXNVIKVG 216

RESULT 18  
US-10-282-122A-64609  
; Sequence 64609, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangsu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John  
; APPLICANT: Carr, Grant  
; APPLICANT: Yamamoto, Robert  
; APPLICANT: Forsyth, R.  
; APPLICANT: Xu, H.  
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
; FILE REFERENCE: ELITRA.034A  
; CURRENT APPLICATION NUMBER: US/10/282,122A  
; CURRENT FILING DATE: 2003-02-20  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR FILING DATE: 2000-09-06

GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 229387  
; LENGTH: 226  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)-(226)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_140794C.1.pep  
US-10-425-115-229387

Query Match 8.1%; Score 188; DB 15; Length 639;  
Best Local Similarity 22.9%; Pred. No. 0.00015;  
Matches 96; Conservative 33; Mismatches 144; Indels 146; Gaps 17;  
QY 7 NNTSSSPGLFQSGDNGLGGHNANSALGQOPIDRQTIEQMAQLLAELLKLLSPQSGNA 66  
DB 159 NGNGGSGAGLWGGGSGGQGGAGANGAAGP-----GK 191  
QY 67 ATGAGGNDQT---TCVGNAGGLNGRKTAGTTTQSDSQNMLSEMNGNLDQAITPDGQ 122  
DB 192 AGSGGNGAGGAGWYGHGGHGGAGNGGNA-TAPGGASAGFDGGAGNG-----GSGGRG 245  
QY 123 GGQIGDNPLLKAMLKLIARMMDQSDQFQPGTGNNNSASGTSSSGSGSPND-----174  
DB 246 GLLFGNG-----GNGSVGGMGGQGTNDTAGDSAGSGGLGNGGNGAGG 290  
QY 175 LSGGKAPSGNSPESGNSVSTPSPPTSPSPPLDFPSPPTAAAGSTPTVTHDPDPVGS 234  
DB 291 LIGNGGGGDSGAGG-----GTDSTQTGVMN-----GASGGSAGIAGNGGDAGL 334  
QY 235 AGTG-----AGNSVAFTSAGANTVLHDTITVKAGOVFDGKGOTFTAGSLGPGQSENQ 289  
DB 335 VNGGAGGNGGNGAAGSALGTT-----IFGSGGVGSGGSGGNGG-----375  
QY 290 KPLFILEDGLSKNVVTMGDDGADGHIHYGDAKIDNLHVTNVCEDAITVKPSAGKSHVE 349  
DB 376 ---WLFSGSAS-----GGNGGQ---GDA-----GTNGFAGFGSGAGSGGSGWG 412  
QY 350 ITNSSFESASDKILQLNADTNLSVDNVKAKDFGTFVRTNGGQOQGN-WDLNLSHISAEDG 407  
DB 413 AVN-----FGPISVQGGFLF--GHGGDGGNGGVDVAGSLSIQFG 449

RESULT 19  
US-10-282-122A-64363  
; Sequence 64363, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangsu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John  
; APPLICANT: Carr, Grant  
; APPLICANT: Yamamoto, Robert  
; APPLICANT: Forsyth, R.  
; APPLICANT: Xu, H.  
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
; FILE REFERENCE: ELITRA.034A

```
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 64363
; LENGTH: 591
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-10-282-122A-64363

Query Match      8.1%; Score 186.5; DB 15; Length 591;
Best Local Similarity 23.1%; Pred. No. 0.00017;
Matches 91; Conservative 31; Mismatches 119; Indels 153; Gaps 15;

QY 6 LNNNTSSPGLFGSGDNLG-----GHNANSALGQPIDRQTTEQMAQLLAEKLSLSP 61
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
116 IGGANGAGPGTQAGGGGLLFGNGGSGCAPQ-----AG 152
QY 62 QSGNAA-----TGAGNDQTTGVGNAG----- 83
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
153 GAGGAAGFFGNGGSGGAGAGGAGGTAGWFFGFGNGGAGGIGVAGINGLGGAGGD 212
QY 84 -----GLNGRKGTAG-----TTQSDSQNMLSENGNGLDQAITPDGQGG 123
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
213 GGNAGFFGNGGSGGAGAGVAVNPGLATPVPAAN-----GGNGLVLGVPGTAGG 267
QY 124 GQIGDNPFLKAMKLIARMMDGSDQFGPGTGNNSASSG-----TSSGSGSPFNDLSG 177
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
268 GADGAN-----GSAIGQAG--GAGGDDGNASTSGGIGIAQTGGAGGA--GGAGG 312
QY 178 GKAPSGNSPSGNYSPVSTFPSPPTSPSTPLDFPSPPTKAAGSTPVTDHPDPVGSAGI 237
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
313 DGAPGGGNG-----GSVEHTGATGSSASGNGATGNGGNGVAGP- 353
QY 238 GAGNSVAFTSAGANQTVLHDTITVKAGOVFDGKQTFTAGSELGDCGSENQKPLFILED 297
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
354 GAGGNGHVSGSVNT-----AGA--GGKGGNGTGGAGCPGCH----- 390
QY 298 GASLKNVTMGDDGADGHLHYGDADKIDNLHVTNNGV 331
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
391 GGSVLSPGVDSGNGGAGGAGGAGVSATDIAGTG 424

RESULT 20
US-10-282-122A-64895
; Sequence 64895, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
```

```
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 64895
; LENGTH: 1381
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-10-282-122A-64895

Query Match      8.0%; Score 185.5; DB 15; Length 1381;
Best Local Similarity 25.0%; Pred. No. 0.00058;
Matches 79; Conservative 25; Mismatches 141; Indels 71; Gaps 11;

QY 19 SGGDNGLGCHNANSALGQ--PIDRQTTEQMAQLLAEKLSLSPQSGNAATGAGNDQT 76
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
779 SGGDGGKGGSGSGGTGGSGAPI-----GGGAGGTGGSGGH 814
QY 77 TGVNAGLNGRKGTAGTTPQSDSQNMLSENGNGLDQAITPDGQGG--GQIGDNPFLKAM 135
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
815 AKGGAGGI--GAQGTITVPNGGN--AGDGGNGNAGA---CGNGSGDPFGNT----- 863
QY 136 LKLARIAMDGSDQFGPGTGNNSASSGTSSTSSGSGSPFNDLSGCKAPS--GNSPSGNYSPVS 194
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
864 -----TSGASGSGGNGNAGTAGSGGAGTGTGLSGGNGGNGGNGGNGGNGG 913
QY 195 TFSPPSTPTSPPLDFPSPPTKAAGSTPVTDHPDPVGSAGIAGNSVAFTSAGANQTV 254
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
914 AHGTVGAQFVPATSLPTPNGGAGGNGGTGNGGAGPAGAPGPTTGGNAGSQGIGGGGNG 973
QY 255 LHDITITKAGVFDG-----KGQTTTAGS-----ELGDDGGSSENQKPLFILED 297
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
974 GGD-----GGKGGDGADAVNVVFMPTBFOAATGTAGSAGDPTGNGGPGTGPSPMVAPPP 1028
QY 298 GASLKNVTMGDDGADG 313
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1029 PTPITVQQGGDGGAG 1044
```

Search completed: January 25, 2005, 12:09:46  
Job time : 62 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2005, 12:02:31 ; Search time 38 Seconds  
(without alignments)  
1131.813 Million cell updates/sec

Title: US-09-879-248-6  
Perfect score: 2310  
Sequence: 1 MSLLTNNNTSSPGLFQSG.....LGDVENHYKVPMSANLKVAE 447

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2310	100.0	447	2 T18447	HrpW protein - Erw
2	353.5	15.3	221	2 A70045	pectate lyase homo
3	233.5	10.1	219	2 S68364	pectate lyase (EC
4	207	9.0	1660	2 A70869	hypothetical glyci
5	199.5	8.6	1306	2 A70934	hypothetical glyci
6	193	8.4	694	2 F70868	hypothetical glyci
7	188	8.1	639	2 D70931	hypothetical glyci
8	186.5	8.1	591	2 B70523	hypothetical glyci
9	186	8.1	1329	2 E70917	hypothetical glyci
10	185.5	8.0	1381	2 E70806	hypothetical glyci
11	185.5	8.0	1489	2 D70807	pectate lyase (EC
12	183.5	7.9	242	2 A45724	hypothetical glyci
13	183	7.9	588	2 F70971	hypothetical glyci
14	182	7.9	606	2 H70816	hypothetical glyci
15	182	7.9	721	2 C70874	hypothetical glyci
16	181.5	7.9	484	2 G70846	hypothetical glyci
17	180.5	7.8	1079	2 B70807	hypothetical glyci
18	179	7.7	730	2 F96559	hypothetical prote
19	179	7.7	1538	2 H70846	hypothetical glyci
20	179	7.7	3016	2 S77300	hypothetical prote
21	178	7.7	741	2 G70917	hypothetical glyci
22	177.5	7.7	1901	2 F70806	hypothetical glyci
23	177.5	7.7	2232	2 T34434	hypothetical prote
24	176.5	7.6	562	2 B70953	hypothetical glyci
25	175	7.6	645	2 F70825	probable PPE prote
26	174	7.5	615	2 H70589	hypothetical glyci
27	172.5	7.5	584	2 G70804	hypothetical glyci
28	172	7.4	3507	2 T34513	hypothetical prote
29	171.5	7.4	667	2 A70893	hypothetical glyci

30	171.5	7.4	1778	2	T50074	probable nucleopor
31	171	7.4	463	2	B70893	hypothetical glyci
32	169.5	7.3	914	2	H70987	hypothetical glyci
33	169	7.3	525	2	D70878	hypothetical glyci
34	168.5	7.3	398	1	OZZQAS	circumsporozoite p
35	167.5	7.3	1011	2	F70620	hypothetical glyci
36	167.5	7.3	2174	2	E95965	hypothetical glyci
37	166.5	7.2	532	2	F70580	hypothetical glyci
38	166.5	7.2	714	2	A70807	hypothetical glyci
39	166	7.2	461	2	F70571	hypothetical glyci
40	166	7.2	603	2	A70770	hypothetical glyci
41	166	7.2	1207	2	T23754	hypothetical prote
42	165.5	7.2	344	2	S41707	PopAI protein - Ps
43	165	7.1	591	2	S74999	iron-regulated pro
44	165	7.1	749	2	A70812	hypothetical glyci
45	165	7.1	959	2	B44402	nuclear pore compl
46	164	7.1	767	2	E70895	hypothetical glyci
47	164	7.1	839	2	H90577	lipoprotein vsa1 f
48	163	7.1	618	2	A70989	hypothetical glyci
49	162	7.0	434	2	E70768	hypothetical glyci
50	161.5	7.0	576	2	A70900	hypothetical glyci
51	161.5	7.0	837	2	E70835	hypothetical glyci
52	161	7.0	787	2	T00798	hypothetical prote
53	160.5	6.9	314	2	F70766	hypothetical prote
54	160.5	6.9	853	2	A70896	hypothetical glyci
55	160.5	6.9	937	2	S58135	hypally regulated
56	160.5	6.9	957	2	D70835	hypothetical glyci
57	159.5	6.9	778	2	F70963	hypothetical glyci
58	159.5	6.9	786	2	T16509	hypothetical prote
59	159	6.9	286	2	C61615	sericin MG-2 - gre
60	159	6.9	1341	2	H98323	hypothetical prote
61	158.5	6.9	1275	2	T33369	hypothetical prote
62	157.5	6.8	2329	2	T28125	hypothetical glyci
63	157	6.8	491	2	D70916	hypothetical glyci
64	157	6.8	543	2	F70726	hypothetical glyci
65	156.5	6.8	496	2	H70839	hypothetical glyci
66	156.5	6.8	1428	2	T08852	lustrin A - Califo
67	155.5	6.7	572	2	T08509	trbl protein - Ent
68	155	6.7	1113	2	S28925	nuclear pore compl
69	155	6.7	1672	2	C81675	polymorphic membra
70	155	6.7	2249	2	A41477	190K surface antig
71	155	6.7	3198	2	A43426	collagen alpha 2 f
72	154.5	6.7	354	2	B70663	probable PPE prote
73	154.5	6.7	582	2	F70675	probable PPE prote
74	154	6.7	403	2	T08471	harpin - Erwinia a
75	154	6.7	430	2	S52700	NUP42 protein - ye
76	153.5	6.6	590	2	E70946	probable PPE prote
77	153	6.6	788	2	JS0747	regulatory protein
78	153	6.6	1049	1	CGE075	collagen alpha 1(I
79	153	6.6	1770	2	A71517	hypothetical prote
80	152.5	6.6	552	2	D70604	probable PPE prote
81	152.5	6.6	683	2	A82704	1,4-beta-cellobios
82	152.5	6.6	923	2	E70820	hypothetical glyci
83	152.5	6.6	2020	2	C48399	ABC-type transport
84	152.5	6.6	3190	2	T13828	CREB-binding prote
85	152	6.6	1560	2	T02885	peroxisome prolife
86	151	6.5	256	2	A70514	hypothetical glyci
87	151	6.5	395	2	A41156	circumsporozoite p
88	151	6.5	1190	2	A82615	surface protein XF
89	151	6.5	1844	2	T51890	related to Nup98-N
90	151	6.5	2554	2	AB3528	extracellular seri
91	150.5	6.5	598	2	T38403	probable nucleopor
92	150.5	6.5	1414	1	S23809	collagen alpha 2(I
93	150	6.5	515	2	H70663	hypothetical glyci
94	150	6.5	1070	2	A10484	probable autotrans
95	150	6.5	1407	2	B72078	polymorphic outer
96	149.5	6.5	2340	2	B71704	cell surface antig
97	149	6.5	605	2	T33913	hypothetical prote
98	149	6.5	681	2	AB2155	hypothetical prote
99	148.5	6.4	1028	2	T03516	probable outer mem
100	148	6.4	387	2	C41156	circumsporozoite p

## ALIGNMENTS

RESULT 1  
 T18447  
 HrpW protein - Erwinia amylovora  
 C:Species: Erwinia amylovora  
 C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
 C:Accession: T18447  
 R:Gaudriault, S.  
 submitted to the EMBL Data Library, May 1998  
 A:Reference number: Z18936  
 A:Accession: T18447  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-447 <GAU>  
 A:Cross-references: UNIPROT:O54508; EMBL:Y13831; NID:e1293450; PID:CAA741  
 A:Experimental source: strain CFBP1430; specific host Pomoideae  
 C:Genetics:  
 A:Note: hrpW

Query Match 100.0%; Score 2310; DB 2; Length 447;  
 Best Local Similarity 100.0%; Pred. No. 3.3e-119; Indels 0; Gaps 0;  
 Matches 447; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSILTLNNNTSSPGIFOSGGDGLGHNANSALGQQPIDRQTIEQMAQLLAELKSLLS 60  
 DB 1 MSILTLNNNTSSPGIFOSGGDGLGHNANSALGQQPIDRQTIEQMAQLLAELKSLLS 60  
 QY 61 POSGNAATGAGGNDQTTGVNAGGLNGKRTAGTTTQSDSONMLSENGNGLDQAITPDG 120  
 DB 61 POSGNAATGAGGNDQTTGVNAGGLNGKRTAGTTTQSDSONMLSENGNGLDQAITPDG 120  
 QY 121 QGGGIGDNPPLKAMKLIARMDQSDQFQPGGTGNNSSAGTSSSGSPFNDLSGGA 180  
 DB 121 QGGGIGDNPPLKAMKLIARMDQSDQFQPGGTGNNSSAGTSSSGSPFNDLSGGA 180  
 QY 181 PSNGSPSGNYSPVSTFSPSTPTSPDPPSPPTKAAGSTPTVTHPPVPVSGAGIGAG 240  
 DB 181 PSNGSPSGNYSPVSTFSPSTPTSPDPPSPPTKAAGSTPTVTHPPVPVSGAGIGAG 240  
 QY 241 NSVAFTSAGANQTVLHDTITVKAQVDFGKGTFTAGSELGQSQENQKPLFILEDGAS 300  
 DB 241 NSVAFTSAGANQTVLHDTITVKAQVDFGKGTFTAGSELGQSQENQKPLFILEDGAS 300  
 QY 301 LKNTMTGDDGADGHLHYGDAKIDNLHVTNVGEDAITVKPNSAGKSHVEITNSSFEHSD 360  
 DB 301 LKNTMTGDDGADGHLHYGDAKIDNLHVTNVGEDAITVKPNSAGKSHVEITNSSFEHSD 360  
 QY 361 KILQNLADTNLSVDNVKAKDFGTVRTNGGQGNWDLNLSHISAEDGKFSFKVSDSEGLN 420  
 DB 361 KILQNLADTNLSVDNVKAKDFGTVRTNGGQGNWDLNLSHISAEDGKFSFKVSDSEGLN 420  
 QY 421 VNTSDISLGDVENHYKVPMSANLKVAE 447  
 DB 421 VNTSDISLGDVENHYKVPMSANLKVAE 447

RESULT 2  
 A70045  
 pectate lyase homolog yypA - Bacillus subtilis  
 C:Species: Bacillus subtilis  
 C>Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 09-Jul-2004  
 C:Accession: A70045  
 R:Kunst, F.; Ogasawara, N.; Moser, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertez  
 C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho  
 A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrazi, E.  
 Nature 390, 249-256, 1997  
 A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galle  
 tech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.  
 Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,  
 A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maue  
 Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parrio, V.; Pohl, T.M.; Portet

Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,  
 A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seror,  
 akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,  
 T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K  
 A:Authors: Yoshikawa, H.P.; Zumstein, E.; Yoshikawa, H.; Danchin, A.  
 A>Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.  
 A:Reference number: A69580; MUID:98044033; PMID:9384377  
 A:Accession: A70045  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-221 <KUN>  
 A:Cross-references: UNIPROT:O34310; GB:Z99121; GB:AL009126; NID:92635827; PIDN:CAB15500  
 A:Experimental source: strain 168  
 C:Genetics:  
 A:Gene: yypA

Query Match 15.3%; Score 353.5; DB 2; Length 221;  
 Best Local Similarity 41.8%; Pred. No. 1e-12; Indels 5; Gaps 1;  
 Matches 76; Conservative 31; Mismatches 70; Indels 5; Gaps 1;

QY 250 ANQTVLHDTITVKAQVDFGKGTFTAGSELGQSQENQKPLFILEDGASLKNTMTGDD 309  
 DB 27 AADKVHETIIVPKNTTYDGKQRFVAGKELGDSQSENQDVPFRVEDGATLKVVVLGAP 86  
 QY 310 GADGHLHYGDAKIDNLHVTNVGEDAITVKPNSAGKSHVEITNSSFEHSDKILQNLADT 369  
 DB 87 AADGVHTYGNVNIQNVKEDVGEDALTVK-----KEGKVTIDGSAQKASDKIFQINKAS 141  
 QY 370 NLSVDNVKAKDFGTVRTNGGQGNWDLNLSHISAEDGKFSFKVSDSEGLNVTNSDLSIG 429  
 DB 142 TPTVKNFTADNGKFRQLGGSTFHVVDIICKTITNKEALFRDSTKSTVMTNTRY 201  
 QY 430 DV 431  
 DB 202 NV 203

RESULT 3  
 S68364  
 pectate lyase (EC 4.2.2.2) C - fungus (Fusarium solani)  
 C:Species: Fusarium solani  
 C>Date: 15-Feb-1997 #sequence\_revision 13-Mar-1997 #text\_change 09-Jul-2004  
 C:Accession: S68364  
 R:Guo, W.; Gonzalez-Candela, L.; Kolattukudy, P.E.  
 Arch. Biochem. Biophys. 323, 352-360, 1995  
 A>Title: Cloning of a new pectate lyase gene pelC from Fusarium solani f. sp. pisi (Nectr  
 A:Reference number: S68364; MUID:96063610; PMID:7487098  
 A:Accession: S68364  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-219 <GUO>  
 A:Cross-references: UNIPROT:Q00843; EMBL:U13049  
 C:Genetics:  
 A:Gene: pelC  
 A:Introns: 52/1; 102/3  
 C:Keywords: carbon-oxygen lyase

Query Match 10.1%; Score 233.5; DB 2; Length 219;  
 Best Local Similarity 30.1%; Pred. No. 3.6e-06;  
 Matches 63; Conservative 39; Mismatches 72; Indels 35; Gaps 6;

QY 220 GGSTPTVTHDPPVGSAGIGAGNSVAFTSAGANQTVLHDTITVKAQVDFGKGTFTAGSE 279  
 DB 8 GGVPKPTDH-----ISNS-----KVIEKAGQVYDGKWKAYDRGSG 43  
 QY 280 LGDG-QGSENQKPLFILEDGASLKNTMTGDDGADGHLHYGDAKIDNLHVTNVGEDAITVK 338  
 DB 44 ACKQNEGGDKDAVFLHNEGATLKNVIIKDSGEVHCXGCHTLEFVWFPEVCEDAISIK 103  
 QY 339 PNSAGKSHVEITNSSFEHSDKILQNLADTNLSVDNVKAKDFGTVRT--NGGQGNWD 396  
 DB 104 EDKAGKESW--IIGGAYHASDKVVGHGCGTWNINFFVEDYKLYRSCGNSCKQCKEN 161

QY 397 LNLHLSAEDKESFVKSDSEGLNNTSD 425  
 Db 162 VYIEGVTAKNG-----GELAGINANYGD 184

RESULT 4

A70869  
 Hypothetical glycine-rich protein Rv2490c - Mycobacterium tuberculosis (strain H37RV)  
 C;Species: Mycobacterium tuberculosis  
 C;Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 09-Jul-2004  
 C;Accession: A70869  
 R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
 Nature 393, 537-544, 1998  
 A;Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
 A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
 A;Reference number: A70500; MUID:98295987; PMID:9634230  
 A;Accession: A70869  
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A;Molecule type: DNA  
 A;Residues: 1-1660 <COL>  
 A;Cross-references: UNIPROT:O53215; GB:AL021246; GB:AL123456; NID:g3261507; PIDN:CAA1606  
 A;Experimental source: strain H37RV  
 C;Genetics:  
 C;Superfamily: collagen alpha 1(IV) chain

Query Match 9.0%; Score 207; DB 2; Length 1660;  
 Best Local Similarity 28.1%; Pred. No. 0.0012;  
 Matches 91; Conservative 19; Mismatches 130; Indels 84; Gaps 15;  
 QY 7 NNTSSPGLFQSGDNLGHNANSALGQOPIRDTQIEQMAQLLAEKLSLSP--QSG 64  
 Db 476 NGCAGGNGGLV--GDGAGCGHGGGAAG-----AGYADMTAIFLGSSTGPGEDGG 523  
 QY 65 N-----AATGAGG-NDQTTGVGNAGLNGRGTAGTTTQSDSQNMLSEMNGNLDOA--IT 117  
 Db 524 NGCAGGAGGAGGAHAGDGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 593  
 QY 118 PDGQGGQIGDNLKLLKMLKLIARMDGQSDQFGQPGTGTNNSSASSTGSSGG--SPFND 174  
 Db 584 GDGAGAGGADAPAGRA-----GSGVGGDGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 633  
 QY 175 LSGGKAPSGNSP--SGNYSVPSTFSPSTPTSPDLPFPSSPTKAAAGSTPVTDPDPVG 233  
 Db 634 GDGAGG 674  
 QY 234 SAGIAGNSVAFTSAGANQTVLHDTITVKAGQVFDGKGQTTAGSELGDCGQSENOKPLF 293  
 Db 675 NGGKG--GNGADATVAGAN-----GGKGGAGGNGGLVGDGGAG----- 710  
 QY 294 ILEDG---ASLKNVTMGDDGADG 313  
 Db 711 --GDGSGGAAGANGNVGDDGADG 732

RESULT 5

A70934  
 Hypothetical glycine-rich protein Rv0578c - Mycobacterium tuberculosis (strain H37RV)  
 C;Species: Mycobacterium tuberculosis  
 C;Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 09-Jul-2004  
 C;Accession: A70934  
 R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
 Nature 393, 537-544, 1998  
 A;Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
 A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
 A;Reference number: A70500; MUID:98295987; PMID:9634230  
 A;Accession: A70934  
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A;Molecule type: DNA

A;Residues: 1-1306 <COL>  
 A;Cross-references: UNIPROT:O53775; GB:AL021942; GB:AL123456; NID:g3242298; PIDN:CAA1744;  
 A;Experimental source: strain H37RV  
 C;Genetics:  
 A;Gene: Rv0578c  
 C;Superfamily: collagen alpha 1(IV) chain

Query Match 8.6%; Score 199.5; DB 2; Length 1306;  
 Best Local Similarity 26.0%; Pred. No. 0.0023;  
 Matches 88; Conservative 16; Mismatches 122; Indels 113; Gaps 11;  
 QY 8 NNTSSPGLFQSGDNLGHNANSALGQOPIRDTQIEQMAQLLAEKLSLSPQSGNAA 67  
 Db 754 NGVAGSQGGAGGDDGTGCVGNGRGRGIDGADGAT----- 789  
 QY 68 TGAGGNDQTTGVGNAGLNGRGTAGTTTQSDSQNMLSEMNGNLDOAITPDGQGGQIG 127  
 Db 790 --AGARGQGGAGGAGGKGRGTGTP-----CGAGPAGTTGSGAGGNG 832  
 QY 128 DNPILLKMLKLIARMDGQSDQFGQPGTGTNNSSASSTGSSGGSPFNDLSGKA--PSGNSP 186  
 Db 833 -----GSGGTGDPDGGGANGSVFTNNGIGGNGGNGGAGPSGAGG 875  
 QY 187 SGNYSVPSTFSPSTPTSPDLP-----DFPSSPTKAAAG-----G 221  
 Db 876 SGGAG--STFG-----ATGSSSIHVNGGNGGNGGNGGNGGNGGNGGNGSLRG 929  
 QY 222 STPVTDPDPVGSAGIGAGNSVAFTSAGANQTVLHDTITVKAGQVFDGKGQTTAGSELG 281  
 Db 930 SGGAGGAGGNGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 973  
 QY 282 DGGQSENOKPLFL-----EDGASLKNVTMGDDGADG 313  
 Db 974 DGGTGDGNGPAGTATGSGRGGGDDGCVGGGSGVAGDGDG 1012

RESULT 6

A70868  
 Hypothetical glycine-rich protein Rv2487c - Mycobacterium tuberculosis (strain H37RV)  
 C;Species: Mycobacterium tuberculosis  
 C;Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 09-Jul-2004  
 C;Accession: A70868  
 R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
 Nature 393, 537-544, 1998  
 A;Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
 A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
 A;Reference number: A70500; MUID:98295987; PMID:9634230  
 A;Accession: A70868  
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A;Molecule type: DNA  
 A;Residues: 1-694 <COL>  
 A;Cross-references: UNIPROT:O53212; GB:AL021246; GB:AL123456; NID:g3261507; PIDN:CAA1606/  
 A;Experimental source: strain H37RV  
 C;Genetics:  
 C;Superfamily: uncharacterized glycine-rich protein, PE motif containing

Query Match 8.4%; Score 193; DB 2; Length 694;  
 Best Local Similarity 25.4%; Pred. No. 0.0024;  
 Matches 86; Conservative 28; Mismatches 140; Indels 84; Gaps 14;  
 QY 18 QSGDNLGAGH--NANSALGQOPIRDTQIEQMAQLLAEKLSLSPQSGNAAATGAGGNDQT 76  
 Db 256 QTGGDGGTGGGTAGTPTGGTGGDGT-----ATAGSKATGAGGDDGT 300  
 QY 77 TGVGNAGLNGRGTAGTTTQSDSQNMLSEMNGNLDOAITPDGQGGQIGDNPILLKML 136  
 Db 301 AAAGCGGNGGDDGVA-----QGDIAAFGDDGNGSDGVAAGSGGSGGAGGG-----AFV 352  
 QY 137 KLIARMDGQSDQFGQPGTGTNNSSASSTGSSGGSPFNDLSGKA--APSGNSPSGNSPV 193

Db 353 HIATATCGSGGFG--GNGAASAGSADGAGGAGGAGGGLLPDGGNGGAGGAGGI 410  
 Qy 194 STSPBPETPTSPPLDPPSSPTKAAAGSTPVT--DHPDP-----VGSAG 236  
 Db 411 GG-----DGTATGPGSGGNAGTARFDPPEAFDPVVGKGGDGGKGGSG 456  
 Qy 237 IGAGNSVAFTSAGANQTVLHDTITVKAGQVDFGKQQTETAGSELGDGGQSENQKPLFILE 296  
 Db 457 LGVGGAGGTGGAGNG-----GAGGLFPGNGN--GGNAGAG-----D 493  
 Qy 297 DGASLKNVTMGDDGADGIHLYGDAKIDNL-HVTNVGED 333  
 Db 494 CGAGVAGGVGNGGGGTATFHEDPVAGVWAVGVGGD 531

RESULT 7  
 D70931  
 hypothetical glycine-rich protein Rv1803c - Mycobacterium tuberculosis (strain H37RV)  
 C:Species: Mycobacterium tuberculosis  
 C>Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 31-Mar-2003  
 C:Accession: D70931  
 R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.; Nature 393, 537-544, 1998  
 A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
 A>Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
 A:Reference number: A70500; MUID:98295987; PMID:9634230  
 A:Accession: D70931  
 A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-639 <COL>  
 A:Cross-references: GB:AL02021; GB:AL123456; NID:g3250699; PIDN:CAA17724.1; PID:e125461  
 A:Experimental source: strain H37RV  
 C:Genetics:  
 A:Gene: Rv1803c  
 C:Superfamily: uncharacterized glycine-rich protein, PE motif containing

Query Match 8.1%; Score 188; DB 2; Length 639;  
 Best Local Similarity 22.9%; Pred. No. 0.004;  
 Matches 96; Conservative 33; Mismatches 144; Indels 146; Gaps 17;  
 Qy 7 NNNTSSPGLFQSGDNGLGGHNANALGQPIDRQTIETQMAQLLAELKSLSPSGNA 66  
 Db 159 NCGSGSAGLWGSNGGQGGAGANGAAGP-----GK 191  
 Qy 67 ATGAGNDOT-----TGVGNAGLNGRKGTAGTTTQSDSQNMLSEMNGLDQAITPDGQG 122  
 Db 192 AGSGGNGGAGGWIYGHGCGAGGNGNA-TAPGASAFDGGAGNG-----GSGRG 245  
 Qy 123 GQIGDNPLLKAMLKLIARMMDGSDQFGQPGTGNNSASGTSSSGSPFND----- 174  
 Db 246 GLLFNG-----GNGSVGGMGGQGTNDTAGDSAGSGGLGNGGNGAQQGW 290  
 Qy 175 LSGKAPSGNSPQNSPVSTFTSPPTPTPTPLDPPSPPTKAAAGSTPVTTHDPPVGS 234  
 Db 291 LIGNGGQGGSGAGG-----GTDSTQTQVMN-----GASGSGAGIAGNGGDA 334  
 Qy 235 AGIG-----AGNSVAFTSAGANQTVLHDTITVKAGQVDFGKQQTETAGSELGDGGQSENQ 289  
 Db 335 VNGGAGCGNGGAGSALGTT-----IFGSGGVGSGGSDGGNGG----- 375  
 Qy 290 KPFILEDGLASLKNVTMGDDGADGIHLYGDAKIDNLHVTNVGEDATVTKPNSAGKSHVE 349  
 Db 376 ---WLFQSGAS-----CGNGGGG---GDA-----GTNGAFGFGSGGGGVMVG 412  
 Qy 350 ITNSSFEHASDKTLQLNADTNLSVDNVKADPRTFTVTVNGGQGN-WDLNLSHTSABDG 407  
 Db 413 AVN-----FQPISVQGGFLF--GHGGDGGGSDVAGSLSIQFG 449

RESULT 8  
 B70523  
 hypothetical glycine-rich protein Rv1450c - Mycobacterium tuberculosis (strain H37RV)  
 C:Species: Mycobacterium tuberculosis  
 C>Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 20-Jun-2000  
 C:Accession: E70917  
 R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.; Nature 393, 537-544, 1998  
 A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
 A>Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
 A:Reference number: A70500; MUID:98295987; PMID:9634230  
 A:Accession: E70917  
 A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-1329 <COL>  
 A:Cross-references: GB:Z95844; GB:AL123456; NID:g3250713; PIDN:CAB09271.1; PID:g2131046  
 A:Experimental source: strain H37RV

hypothetical glycine-rich protein Rv0297 - Mycobacterium tuberculosis (strain H37RV)  
 C:Species: Mycobacterium tuberculosis  
 C>Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 31-Mar-2003  
 C:Accession: B70523  
 R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.; Nature 393, 537-544, 1998  
 A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
 A>Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
 A:Reference number: A70500; MUID:98295987; PMID:9634230  
 A:Accession: B70523  
 A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-591 <COL>  
 A:Cross-references: GB:Z96800; GB:AL123456; NID:g3261800; PIDN:CAB09596.1; PID:e321655, I  
 A:Experimental source: strain H37RV  
 C:Genetics:  
 A:Gene: Rv0297  
 C:Superfamily: uncharacterized glycine-rich protein, PE motif containing

Query Match 8.1%; Score 186.5; DB 2; Length 591;  
 Best Local Similarity 23.1%; Pred. No. 0.0044;  
 Matches 91; Conservative 31; Mismatches 119; Indels 153; Gaps 15;  
 Qy 6 LNNTSSPGLFQSGDNGL-----GHNANALGQPIDRQTIETQMAQLLAELKSLSP 61  
 Db 116 IINGANGAPGTQAGDGGLLFNGGSGGAPGQ-----AG 152  
 Qy 62 QSGNAA-----TGAGNDQTTGVGNAG----- 83  
 Db 153 GAGGAGFFGNGGCGDGGAGGAGGTAGTFFGFGGAGGIVAGINGLGGAGD 212  
 Qy 84 -----GLNGRKGTAG-----TTPOSQNMSEMNGLDQAITPDGQG 123  
 Db 213 CGNAGFFGNGGNGGAGAGVNAVNPGLATPTVPAAN-----GNGMLNVGPVTAGG 267  
 Qy 124 GQIGDNPLLKAMLKLIARMMDGSDQFGQPGTGNNSASG-----TSSSGSGSPFNDLSG 177  
 Db 268 GADGAN-----GSAIGQAG--GAGDGGNASTSGGIGIAQTGGAGGA--CGAGG 312  
 Qy 178 GKAPSGNSPQNSPVSTFTSPPTPTPTPLDPPSPPTKAAAGSTPVTTHDPPVGSAGI 237  
 Db 313 DGAPFGNGNG-----GSVEHTCATGSSAGSGNGATGNGGNGVGAPO- 353  
 Qy 238 GAGNSVAFTSAGANQTVLHDTITVKAGQVDFGKQQTETAGSELGDGGQSENQKPLFILED 297  
 Db 354 GAGGNGGHVSGSVNT-----AGA--GGKGGNGGTGGAGGPGGH----- 390  
 Qy 298 GASLKNVTMGDDGADGIHLYGDAKIDNLHVTNVG 331  
 Db 391 GGSVLSPVGDGNGGAGGAGGAGGAGVSATDIAGTG 424

RESULT 9  
 E70917  
 hypothetical glycine-rich protein Rv1450c - Mycobacterium tuberculosis (strain H37RV)  
 C:Species: Mycobacterium tuberculosis  
 C>Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 20-Jun-2000  
 C:Accession: E70917  
 R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.; Nature 393, 537-544, 1998  
 A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
 A>Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
 A:Reference number: A70500; MUID:98295987; PMID:9634230  
 A:Accession: E70917  
 A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-1329 <COL>  
 A:Cross-references: GB:Z95844; GB:AL123456; NID:g3250713; PIDN:CAB09271.1; PID:g2131046  
 A:Experimental source: strain H37RV







A>Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
A/Reference number: A70500; MUID:98295987; PMID:9634230  
A/Accession: C70974  
A/Status: preliminary; nucleic acid sequence not shown; translation not shown  
A/Molecule type: DNA  
A/Residues: 1-731 <COL>  
A/Cross-references: GB:AL009198; GB:AL123456; NID:g3242262; PIDN:CAA15773.1; PID:g266166  
A/Experimental source: strain H37Rv  
C/Genetics:  
A/Gene: Rv3388  
C/Superfamily: elastin related uncharacterized glycine-rich protein, PE motif containing

Query Match 7.9%; Score 182; DB 2; Length 731;  
Best Local Similarity 25.3%; Pred. No. 0.01;  
Matches 89; Conservative 26; Mismatches 127; Indels 110; Gaps 15;

QY 7 NNNTSSPGLFQSQGDNLGCHNANSALGQQPIDRTQIEQAQLLAELLKSLSPSQSNA 66  
DB |||:::||||| |  
399 NNTSAGTGGVAGSGGTGNGAG-----LIGAGHGG 428

QY 67 ATGAGGNDDTTGVGNAGLGRKGTAGTTTPQ-----SDSN-----MLSEMGNNGLDQAIT 117  
DB |||:::||||| |

DB 429 AGGAGGN-QTGVCVN-GGAGGCGAGGAGGQLYNGGSDGGGAGGANIAGNGSDGGAA 486  
|||:::||||| |

QY 118 PDGQGGGQIGDNPLLKAMLLKIARMDDQSDQFGPGCTGNNSAS-----SGTSSSGGSP 171  
DB |||:::||||| |

DB 487 GHGAGGS-----ARLI-GAGGHGGDGAGGNTAGRADAATAGTCGGGNG 531  
|||:::||||| |

QY 172 FNDLSGCKAPSGNSPSNYSPVSTFTSPPTPTPTPTPLDFPSSPTKAAGGSTPTVDHPDP 231  
DB |||:::||||| |

DB 532 GN-----GCLLSGNAGAGGHHGAGGSATTATTG-----TPPTGATGNN----- 570  
|||:::||||| |

QY 232 VGSAGIGAGNSVAFTSAGANQTVLHDHTITVKAGVFQKGOTFTAGSELCDGCQSENQKP 291  
DB |||:::||||| |

DB 571 -GGNG-GAGGTAGTGTSG-----GICGNGAGGTGNGAVA----- 604  
|||:::||||| |

QY 292 LFILEDGASLNVTMGBDDGADGHLHYGDAKIDLHVTVNGEDAITVKPNASG 343  
DB |||:::||||| |

DB 605 LSVGSTGLGNGSGSLGCGGSLFCNGGAGGVGATG-GNGGSGIGCPASVG 655  
|||:::||||| |

RESULT 16  
G70846  
hypothetical glycine-rich protein Rv3344c - Mycobacterium tuberculosis (strain H37RV)  
C/Species: Mycobacterium tuberculosis  
C/Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 20-Jun-2000  
C/Accession: G70846  
R/Cole, S.F.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feldwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998  
A/Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrall, B.G.  
A/Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
A/Reference number: A70500; MUID:98295987; PMID:9634230  
A/Accession: G70846  
A/Status: preliminary;  
A/Molecule type: DNA  
A/Residues: 1-484 <COL>  
A/Cross-references: GB:AL021841; GB:AL123456; NID:g3261517; PIDN:CAA17116.1; PID:g289425  
A/Experimental source: strain H37Rv  
C/Genetics:  
A/Gene: Rv3344c  
C/Superfamily: Phaseolus glycine-rich cell wall protein 1.8

Query Match 7.9%; Score 181.5; DB 2; Length 484;  
Best Local Similarity 26.2%; Pred. No. 0.0065;  
Matches 90; Conservative 24; Mismatches 108; Indels 121; Gaps 19;

QY 14 PGLFQSQGDNLGCHNANSALGQQPIDRTQIEQAQLLAELLKSLSPSQSNAAT---G 69  
DB PG--STGGAGKGGAGANG-----SSNGNAGNGNAGNHGG 99  
|||:::||||| |

QY 70 AGGNDOTTGVGNAGLGRKGTAGTTTPQSDSNMLSEMGNGLDQAITPDGQG----- 122  
|||:::||||| |

```

Db      100 AGGSDTCGAGAGGGGGTGTGS-----GSGGGAGGNG-----GNGGAGGTGV 146
Qy      123 --GGQIGDNPLLLKAMKLIIARMWGDQSDFG---QPCTGNNSASSGTSSSGGSPFNDL-S 176
Db      147 VLGGKGDDG-----GNGDHGPATNFCGSRGGAGGCGGCGAGGNATCS 191
Qy      177 GKG-APSGNSPSGNYSPVSTFPSPPTPTSPLDPPSSP-----TKAAGGSTPTVTDPHD 230
Db      192 GKGAGGAGGDDGDSFG-----ATSGPASIGVTGAPCGNGGKGACGS-----N 234
Qy      231 PVGSAGIG--AGNSVAFTSAGANQTVLHDTITVKAGOVFDKGQTPTAGSELGGGQSEN 288
Db      235 PNGSGDGGKGGNG-----GAGGN-----GSGIGANS--IVCGSGAGGAGGAGGNGSLSSG 285
Qy      289 QKPLFILEDGASLKNTVMGDDGADGIHLGYCDAKIDNLHVTVNVG 331
Db      286 E-----GGKGDGGH-GDGVGGNSSVTQGG 310

RESULT 17
B70807
hypothetical glycine-rich protein Rv3512 - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
C:Accession: B70807
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A>Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: B70807
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1079 <COL>
A:Cross-references: GB:ALJ23456; NID:g3261554; PIDN:CAA17749.1; PID:g292444
A:Experimental source: strain H37RV
C:Genetics:
C:Gene: Rv3512
C:Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homology;

Query Match          7.8%; Score 180.5; DB 2; Length 1079;
Best Local Similarity 25.3%; Pred. No. 0.019;
Matches      81; Conservative 17; Mismatches 105; Indels 117; Gaps 15;

Qy      7 NNNTSSPGLFSQSGDNGLGHNANSALCQQPIDRTIQEQMAQLLELLSLSPSQNA 66
Db      700 NGGAGTGSG---NGNGSGGNGGNAGMG-----GNS 728
Qy      67 AT----CAGGNDQTTQVGNAGLNCRKCTAGTTPQ----SDSONMLSEMNNGLDQAIT 117
Db      729 GTGSGDGGAGNGGAHTGTGTGGTGGGGTGGDGN-----GCNADNTAN 783
Qy      118 PDGGGGGIGDNPILLKAMLKLIARMDGOSDQPGPCTGNNSASSGTSSSGGSPFNDLSG 177
Db      784 MTAAQGDDGNG-----GDGFGGAGAGGGGLTAGANGTGG-----QG 822
Qy      178 GKAPSNGNSPSGNYSPVSTFPSPPTPTSPLDFFSPSPKAAGGSTPTVTHDPVGSAGI 237
Db      823 GAGDGG---GN-----GAIIGHGLPLTD--DPGNGGT 849
Qy      238 GAGNSVAFTSAGANQTVLHDTITVKAGOVFDKGQTPTAGSEL---GDGQSENQKPLF 293
Db      850 G-GNGGTGGTGGAGTGSJGGGTGGDGN--GGNGGTGGEGEVGGAGGTGGAAGN----- 902
Qy      294 ILEDGASLKNTVMGDDGADG 313
Db      903 --GDGG-----TGGTGGDGD 915

RESULT 18
```

## RESULT 18







GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2005, 11:56:15 ; Search time 64 Seconds  
(without alignments)  
4018.632 Million cell updates/sec

Title: US-09-879-248-6  
Perfect score: 2310  
Sequence: 1 MSLLTNNNTSSPGLFQSG.....LGDVENHYKVPMSANLKVAE 447

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : Uniprot 02.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2310	100.0	447	2	O54508 erwinia amy
2	2277	98.6	447	2	Q9LAW2
3	1916.5	83.0	450	2	O6XDB8
4	1916.5	83.0	450	2	AAQ17046
5	1916.5	83.0	450	2	AAS45453
6	1005	43.5	479	2	O6RK52
7	1005	43.5	479	2	Q6DSC8
8	1005	43.5	479	2	AAS20352
9	705	30.5	138	2	Q79AW7
10	589	25.5	424	2	O8RP12
11	567.5	24.6	441	2	O87327
12	559	24.2	424	2	O87264
13	559	24.2	424	2	Q7C415
14	472.5	20.5	386	2	O849P3
15	423.5	18.3	307	2	O8KKZ2
16	415.5	18.0	380	2	O8XVQ5
17	362.5	15.7	224	2	Q8RHM0
18	354	15.3	303	2	O8PHI5
19	353.5	15.3	221	2	O34310
20	349.5	15.1	222	2	Q9X622
21	346.5	15.0	324	2	O8PBA4
22	341	14.8	222	2	O8LOR5
23	256	11.1	276	2	O829M8
24	226.5	9.8	266	2	Q9EX16
25	216.5	9.4	215	2	Q00843
26	207	9.0	1660	2	Q79FD4
27	207	9.0	1660	2	CAE55496
28	207	9.0	1665	2	Q7D721
29	206	8.9	252	2	Q7S4B7
30	199.5	8.6	1306	2	Q7D9L6
31	199.5	8.6	1306	2	Q7UIQ7 mycobacteri

#### ALIGNMENTS

Q6mx28 mycobacteri  
Caes5300 mycobacte  
Q00845 nectria hae  
Q8t9H1 drosophila  
Q93877 fusarium ox  
Q7d724 mycobacteri  
Caes5495 mycobacte  
Q8x80 drosophila  
Q9nb13 drosophila  
Q9nbW0 drosophila  
Aas64813 drosophila  
Q7kth9 drosophila  
Q9nhQ0 drosophila  
Aas64812 drosophila  
Q9v8E6 drosophila  
Q9n693 drosophila  
Q7tWC4 mycobacteri  
Q6bn56 debaryomyce  
Q7tWC0 mycobacteri  
Q8vJ19 mycobacteri  
Q7t213 mycobacteri  
Q79fJ9 mycobacteri  
Caes5430 mycobacte  
Q8vJw1 mycobacte  
Q8vkn3 mycobacteri  
Q8vU20 mycobacteri  
Q6mx50 mycobacteri  
Caes5268 mycobacte  
Q7u2C0 mycobacteri  
Q9ngw5 drosophila  
Q79fP2 mycobacteri  
Caes5390 mycobacte  
Q7u022 mycobacteri  
Q8vK17 mycobacteri  
P90534 dictyosteli  
Q6mw9 mycobacteri  
Caes5603 mycobacte  
Q6mw6 mycobacteri  
Caes5607 mycobacte  
Q9grA9 drosophila  
Q7tYg8 mycobacteri  
Q89cb5 bradyrhizob  
Q8v1z1 mycobacteri  
Q8v4701 nectria hae  
Q6mwX7 mycobacteri  
Caes5590 mycobacte  
Q8vJ15 mycobacteri  
Q8vJ23 mycobacteri  
Q00851 nectria hae  
Q8vJy9 mycobacteri  
Q79fV3 mycobacteri  
Caes5332 mycobacte  
Q8vKc5 mycobacteri  
Q6mx5 mycobacteri  
Caes5593 mycobacte  
Q7rtD0 plasmodium  
Q6mwy1 mycobacteri  
Caes5586 mycobacte  
Q6fg10 candida gla  
Q7tWk6 mycobacteri  
Q79sh8 neurospora  
Q79uq0 rhodospirell  
Q6mwW7 mycobacteri  
Caes5606 mycobacte  
Q8v1z0 mycobacteri  
Q7tWm2 mycobacteri  
Q9qek6 cynomolgus  
Q7u125 mycobacteri  
Q75f10 leptospira

32 199.5 8.6 1306 2 Q6mx28  
33 199.5 8.6 1306 2 CAE5300  
34 194.5 8.4 233 2 Q00845  
35 194 8.4 1928 2 Q8t9H1  
36 193 8.4 240 2 Q93877  
37 193 8.4 694 2 Q7D724  
38 193 8.4 694 2 CAE5495  
39 191 8.3 555 2 Q8X80  
40 191 8.3 929 2 Q9NBL3  
41 191 8.3 929 2 Q9NBW0  
42 191 8.3 929 2 AAS64813  
43 191 8.3 939 2 Q7KRH9  
44 191 8.3 939 2 Q9NHQ0  
45 191 8.3 939 2 AAS64812  
46 191 8.3 2280 2 Q9V8E6  
47 191 8.3 2302 2 Q9N693  
48 190.5 8.2 1360 2 Q7TWC4  
49 190.5 8.2 1382 2 Q6BN56  
50 190.5 8.2 1382 2 Q7TWC0  
51 189.5 8.2 628 2 Q8VJ19  
52 188 8.1 540 2 Q7T2I3  
53 188 8.1 639 2 Q79FJ9  
54 188 8.1 639 2 CAE55430  
55 188 8.1 650 2 Q8VJW1  
56 187.5 8.1 622 2 Q8VKN3  
57 187 8.1 795 2 Q7U020  
58 186.5 8.1 591 2 Q6MX50  
59 186.5 8.1 591 2 CAE55268  
60 186.5 8.1 606 2 Q7U2C0  
61 186 8.1 929 2 Q9NGW5  
62 186 8.1 1329 2 Q79FP2  
63 186 8.1 1329 2 CAE55390  
64 186 8.1 1408 2 Q7U022  
65 186 8.1 1408 2 Q8VK17  
66 185.5 8.0 809 2 P90534  
67 185.5 8.0 1381 2 Q6MWW9  
68 185.5 8.0 1381 2 CAE55603  
69 185.5 8.0 1489 2 Q6MWW6  
70 185.5 8.0 1489 2 CAE55607  
71 185 8.0 2310 2 Q9GRA9  
72 184.5 8.0 1150 2 Q7TYG8  
73 184 8.0 987 2 Q89CB5  
74 184 8.0 1384 2 Q8V1Z1  
75 183.5 7.9 242 2 Q8V4701  
76 183 7.9 588 2 Q6MWW7  
77 183 7.9 588 2 CAE55590  
78 183 7.9 775 2 Q8VJ15  
79 183 7.9 1507 2 Q8VJ23  
80 182.5 7.9 244 2 Q00851  
81 182.5 7.9 1217 2 Q8V1Y9  
82 182 7.9 606 2 Q79FPV3  
83 182 7.9 606 2 CAE55332  
84 182 7.9 609 2 Q8VKC5  
85 182 7.9 731 2 Q6MWW5  
86 182 7.9 731 2 CAE55593  
87 182 7.9 1991 2 Q7RTD0  
88 181.5 7.9 484 2 Q6MWW1  
89 181.5 7.9 484 2 CAE55586  
90 181.5 7.9 1618 2 Q6FQ10  
91 181 7.8 749 2 Q7TWM6  
92 181 7.8 1376 2 Q7SSH8  
93 181 7.8 3056 2 Q7USQ0  
94 180.5 7.8 1079 2 Q6MWW7  
95 180.5 7.8 1079 2 CAE55606  
96 180.5 7.8 1715 2 Q8V1Z0  
97 180 7.8 626 2 Q7TWM2  
98 179.5 7.8 588 2 Q9QEK6  
99 179.5 7.8 608 2 Q7U125  
100 179.5 7.8 717 2 Q75F10

```

RESULT 1
OS4508
ID O54508 PRELIMINARY; PRT; 447 AA.
AC O54508;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE HrpW protein.
GN Name=hrpW;
OS Erwinia amylovora.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Erwinia.
OX NCBI_TaxID=552;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CFBP1430;
RX MEDLINE=98086111; PubMed=9426142;
RA Gaudriault S., Malandrin L., Paulin J.P., Barny M.A.;
RT "DepA, an essential pathogenicity factor of Erwinia amylovora showing
RT homology with AvrE of Pseudomonas syringae, is secreted via the Hrp
RT secretion pathway in a DspB-dependent way.";
RL Mol. Microbiol. 26:1057-1069(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CFBP1430;
RX MEDLINE=98316710; PubMed=9654138;
RA Gaudriault S., Brisset M.N., Barny M.A.;
RT "HrpW of Erwinia amylovora, a new Hrp-secreted protein.";
RL FEBS Lett. 428:224-228(1998).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Ba321;
RA Kim J.F., Zumoff C.H., Beer S.V.;
RT "HrpW, a new harpin of Erwinia amylovora, is a member of a family of
RT pectate lyases.";
RL Phytopathology 87:0-0(1997).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Ba321;
RX MEDLINE=98422475; PubMed=9748455;
RA Kim J.F., Beer S.V.;
RT "HrpW of Erwinia amylovora, a new harpin that contains a domain
RT homologous to pectate lyases of a distinct class.";
RL J. Bacteriol. 180:5203-5210(1998).
DR EMBL; Y13831; CAA74158.1; -.
DR EMBL; U94513; AAC62314.1; -.
DR PIR; T18447; T18447.
DR HSP; Q9RHW0; IEE6.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0030570; F:pectate lyase activity; IEA.
DR InterPro; IPR011050; Pectin_lyas_like.
DR InterPro; IPR004898; Pect_lyase.
DR Pfam; PF03211; Pectate_lyase; 1.
DR SQ SEQUENCE 447 AA; 45347 MW; 7A744D169D6D9CF3 CRC64;

Query Match 100.0%; Score 2310; DB 2; Length 447;
Best Local Similarity 100.0%; Pred. No. 2.3e-110;
Matches 447; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSILTLNNNTSSPGLFQSGDNGLGHNANSALGQOPIDRQTIQMAQLLAELLSLS 60
DB 1 MSILTLNNNTSSPGLFQSGDNGLGHNANSALGQOPIDRQTIQMAQLLAELLSLS 60

QY 61 PQSGNAATGAGNDQTTGVGNAGLNGRKGTAGTTPQSDSQNMLSEMGNGLDQAITPDG 120
DB 61 PQSGNAATGAGNDQTTGVGNAGLNGRKGTAGTTPQSDSQNMLSEMGNGLDQAITPDG 120

QY 121 QGGGQIGDNPILLKAMLIARMDGQSDQFGQPGTGNNSASSTSSGSGSPNDLSGGKA 180
DB 121 QGGGQIGDNPILLKAMLIARMDGQSDQFGQPGTGNNSASSTSSGSGSPNDLSGGKA 180

QY 181 PSNGSPSGNYSPVSTFSPSTPTSPPLDFFSSPTKAAGGSTPTVDHDPVGSAGIGAG 240
DB 181 PSNGSPSGNYSPVSTFSPSTPTSPPLDFFSSPTKAAGGSTPTVDHDPVGSAGIGAG 240

QY 241 NSVAFTSAGANQTVLHDTITVKAGQVFDGKQOTFTAGSELGDGQGSSEKQKPLFILEDGAS 300
DB 241 NSVAFTSAGANQTVLHDTITVKAGQVFDGKQOTFTAGSELGDGQGSSEKQKPLFILEDGAS 300

QY 301 LKNVTMGDDGADGIGIHLYGDAKIDNLHVTNVGDAITVKPNSAGKSHVEITNSSFFHASD 360
DB 301 LKNVTMGDDGADGIGIHLYGDAKIDNLHVTNVGDAITVKPNSAGKSHVEITNSSFFHASD 360

```

```

181 PSNGSPSGNYSPVSTFSPSTPTSPPLDFFSSPTKAAGGSTPTVDHDPVGSAGIGAG 240
241 NSVAFTSAGANQTVLHDTITVKAGQVFDGKQOTFTAGSELGDGQGSSEKQKPLFILEDGAS 300
241 NSVAFTSAGANQTVLHDTITVKAGQVFDGKQOTFTAGSELGDGQGSSEKQKPLFILEDGAS 300
301 LKNVTMGDDGADGIGIHLYGDAKIDNLHVTNVGDAITVKPNSAGKSHVEITNSSFFHASD 360
301 LKNVTMGDDGADGIGIHLYGDAKIDNLHVTNVGDAITVKPNSAGKSHVEITNSSFFHASD 360
361 KILQLNADTNLSVDNVKAKDFGTFVRTNGGQQNWDNLNLSHISAEDGKTSFVKSDSEGLN 420
361 KILQLNADTNLSVDNVKAKDFGTFVRTNGGQQNWDNLNLSHISAEDGKTSFVKSDSEGLN 420
421 VNTSDISLGDVENHYKVPMSANLKVAE 447
421 VNTSDISLGDVENHYKVPMSANLKVAE 447

RESULT 2
Q9LAW2 PRELIMINARY; PRT; 447 AA.
AC Q9LAW2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Harpin HrpW.
GN Name=hrpW;
OS Erwinia amylovora.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Erwinia.
OX NCBI_TaxID=552;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Ba246;
RA Kim J.F., Laby R.J., Beer S.V.;
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF083620; AAF63402.1; -.
DR HSP; Q9RHW0; IEE6.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0030570; F:pectate lyase activity; IEA.
DR InterPro; IPR011050; Pectin_lyas_like.
DR InterPro; IPR004898; Pect_lyase.
DR Pfam; PF03211; Pectate_lyase; 1.
DR SQ SEQUENCE 447 AA; 45340 MW; 0BBAAE3871EDC2F6 CRC64;

Query Match 98.6%; Score 2277; DB 2; Length 447;
Best Local Similarity 98.7%; Pred. No. 1.1e-108;
Matches 441; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 MSILTLNNNTSSPGLFQSGDNGLGHNANSALGQOPIDRQTIQMAQLLAELLSLS 60
DB 1 MSILTLNNNTSSPGLFQSGDNGLGHNANSALGQOPIDRQTIQMAQLLAELLSLS 60

QY 61 PQSGNAATGAGNDQTTGVGNAGLNGRKGTAGTTPQSDSQNMLSEMGNGLDQAITPDG 120
DB 61 PQSGNAATGAGNDQTTGVGNAGLNGRKGTAGTTPQSDSQNMLSEMGNGLDQAITPDG 120

QY 121 QGGGQIGDNPILLKAMLIARMDGQSDQFGQPGTGNNSASSTSSGSGSPNDLSGGKA 180
DB 121 QGGGQIGDNPILLKAMLIARMDGQSDQFGQPGTGNNSASSTSSGSGSPNDLSGGKA 180

QY 181 PSNGSPSGNYSPVSTFSPSTPTSPPLDFFSSPTKAAGGSTPTVDHDPVGSAGIGAG 240
DB 181 PSNGSPSGNYSPVSTFSPSTPTSPPLDFFSSPTKAAGGSTPTVDHDPVGSAGIGAG 240

QY 241 NSVAFTSAGANQTVLHDTITVKAGQVFDGKQOTFTAGSELGDGQGSSEKQKPLFILEDGAS 300
DB 241 NSVAFTSAGANQTVLHDTITVKAGQVFDGKQOTFTAGSELGDGQGSSEKQKPLFILEDGAS 300
301 LKNVTMGDDGADGIGIHLYGDAKIDNLHVTNVGDAITVKPNSAGKSHVEITNSSFFHASD 360

```



```
Db 301 LKNVTMGDDGADGHIHYGDADKIDNLHVTNVGEDAITVKPNSAGKSHVEITNSSPEHSD 360
QY 361 KILQLNADTNLSVDNVKAKDFGTFVRTNGGQGGNWDNLNLSHISAEDGKFSFVKSDSEGLN 420
Db 361 KILQLNADTNLSVDNVKAKDFGTFVRTNGGQGGNWDNLNLSHISAEDGKFSFVKSDSEGLK 420
QY 421 VNTSDISLGDVENHYKVPMSANLKVAE 447
Db 421 VNTSDISLGDVENHYKVPMSANLKVAE 447

RESULT 3
Q6XDB8
ID Q6XDB8 PRELIMINARY; PRT; 450 AA.
AC Q6XDB8;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE HrpW.
GN Name=hrpW;
OS Erwinia pyrifoliae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Erwinia.
OX NCBI_TaxID=79967;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WT#3;
RA Shrestha R., Baek S.J., Bae H.N., Cho J.M., Park D.H., Hwang I.K.,
RA Hur J.H., Lim C.K.;
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY237642; AAQ17046.1; -.
DR InterPro; IPR011050; Pectin lyase like.
DR Pfam; PF03211; Pectate lyase; 1.
SQ SEQUENCE 450 AA; 45509 MW; 085F90CD322C3A74 CRC64;

Query Match 83.0%; Score 1916.5; DB 2; Length 450;
Best Local Similarity 83.2%; Pred. No. 2.9e-90;
Matches 376; Conservative 23; Mismatches 46; Indels 7; Gaps 2;

QY 1 MSILTLNNTSSPGLFQSGDNGLCGHVANSALGQOPIDRQTIQWQAQLLAELLKLLS 60
Db 1 MSVLTNLISIPSSQGLFKPEDNGLSGQNTNSAQGHPIIDRQTIQWQAQLLGELLKPLLS 60
QY 61 PQSGNAATGAGNDQTTGVGNAGLNGRKGCTAGTTTQSDSQNNMLSEWNGNGLDQAITPDG 120
Db 61 PQADNAA--AGSNDQTNVGNAGLNGRKGCTAGTTTQSDSQNNMLSEWNGNGLDQAITPDG 118
QY 121 QGGGQIGDNPLLKAMLKLIARMMDGSDQFGQPGTGNSSAGTSSGSPFNDS---- 176
Db 119 QGGGQIGDNPLLKAMLKLIARMMDGSDQFGQPGSGNNDASSGSPSAGNYPASNASSGSS 178
QY 177 -GKAPSGNSPSGNSPVSTFSPSTPTSPSLDPPSSPTKAAGSTPTVTDHPDPVGS 235
Db 179 LGSSSLGKASSGSGTPTNSFSPTPTSPSLDPPSSPTKAAGSTPTVTDHADPVGST 238
QY 236 GIGAGNSVAFTSAGANOTVLHDTITVKAGVDFGKGTFTAGSELGDDGGSQENKPLFIL 295
Db 239 GVAGNSVGTTSAGANPTVLHDTIIIVKAGQEFPGKGTFTAGSELGDDGGSQENKPLFIL 298
QY 296 EDGASLKNVTMGDDGADGHIHYGDADKIDNLHVTNVGEDAITVKPNSAGKSHVEITNSSF 355
Db 299 EDGASLKNVTMGDDGADGHIHYGDADKIDNLHVTNVGEDAITVKPNSAGKSHVEITNSSF 358
QY 356 EDGASLKNVTMGDDGADGHIHYGDADKIDNLHVTNVGEDAITVKPNSAGKSHVEITNSSF 355
Db 358 EDGASLKNVTMGDDGADGHIHYGDADKIDNLHVTNVGEDAITVKPNSAGKSHVEITNSSF 358
QY 356 EHASDKILQLNADTNLSVDNVKAKDFGTFVRTNGGQGGNWDNLNLSHISAEDGKFSFVKSD 415
Db 359 EHASDKILQLNADTNLSVDNVKAKDFGTFVRTNGGQGGNWDNLNLSHISAEDGKFSFVKSD 418
QY 416 SEGNAVNTSDISLGDVENHYKVPMSANLKVAE 447
Db 419 SEGNAVNTSDISLGDVENHYKVPMSANLKVAK 450

RESULT 5
AA545453
ID AA545453 PRELIMINARY; PRT; 450 AA.
AC AA545453;
```

```
Db 359 EHASDKILQLNADTNLSVDNVKAKDFGTFVRTNGGQGGNWDNLNLSHISAEDGKFSFVKSD 418
QY 416 SEGNAVNTSDISLGDVENHYKVPMSANLKVAE 447
Db 419 SEGNAVNTSDISLGDVENHYKVPMSANLKVAK 450

RESULT 4
AAQ17046
ID AAQ17046 PRELIMINARY; PRT; 450 AA.
AC AAQ17046;
DT 03-MAR-2004 (TrEMBLrel. 27, Created)
DT 03-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT 03-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE HrpW.
GN HRPW.
OS Erwinia pyrifoliae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Erwinia.
OX NCBI_TaxID=79967;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WT#3;
RA Shrestha R., Baek S.J., Bae H.N., Cho J.M., Park D.H., Hwang I.K.,
RA Hur J.H., Lim C.K.;
RL "Molecular characterization hrp genes cluster of Erwinia pyrifoliae
RL and expression of hrpP encoding elicitor of the hypersensitive
RL response.";
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY237642; AAQ17046.1; -.
SQ SEQUENCE 450 AA; 45509 MW; 085F90CD322C3A74 CRC64;

Query Match 83.0%; Score 1916.5; DB 2; Length 450;
Best Local Similarity 83.2%; Pred. No. 2.9e-90;
Matches 376; Conservative 23; Mismatches 46; Indels 7; Gaps 2;

QY 1 MSILTLNNTSSPGLFQSGDNGLCGHVANSALGQOPIDRQTIQWQAQLLAELLKLLS 60
Db 1 MSVLTNLISIPSSQGLFKPEDNGLSGQNTNSAQGHPIIDRQTIQWQAQLLGELLKPLLS 60
QY 61 PQSGNAATGAGNDQTTGVGNAGLNGRKGCTAGTTTQSDSQNNMLSEWNGNGLDQAITPDG 120
Db 61 PQADNAA--AGSNDQTNVGNAGLNGRKGCTAGTTTQSDSQNNMLSEWNGNGLDQAITPDG 118
QY 121 QGGGQIGDNPLLKAMLKLIARMMDGSDQFGQPGTGNSSAGTSSGSPFNDS---- 176
Db 119 QGGGQIGDNPLLKAMLKLIARMMDGSDQFGQPGSGNNDASSGSPSAGNYPASNASSGSS 178
QY 177 -GKAPSGNSPSGNSPVSTFSPSTPTSPSLDPPSSPTKAAGSTPTVTDHPDPVGS 235
Db 179 LGSSSLGKASSGSGTPTNSFSPTPTSPSLDPPSSPTKAAGSTPTVTDHADPVGST 238
QY 236 GIGAGNSVAFTSAGANOTVLHDTITVKAGVDFGKGTFTAGSELGDDGGSQENKPLFIL 295
Db 239 GVAGNSVGTTSAGANPTVLHDTIIIVKAGQEFPGKGTFTAGSELGDDGGSQENKPLFIL 298
QY 296 EDGASLKNVTMGDDGADGHIHYGDADKIDNLHVTNVGEDAITVKPNSAGKSHVEITNSSF 355
Db 299 EDGASLKNVTMGDDGADGHIHYGDADKIDNLHVTNVGEDAITVKPNSAGKSHVEITNSSF 358
QY 356 EHASDKILQLNADTNLSVDNVKAKDFGTFVRTNGGQGGNWDNLNLSHISAEDGKFSFVKSD 415
Db 359 EHASDKILQLNADTNLSVDNVKAKDFGTFVRTNGGQGGNWDNLNLSHISAEDGKFSFVKSD 418
QY 416 SEGNAVNTSDISLGDVENHYKVPMSANLKVAE 447
Db 419 SEGNAVNTSDISLGDVENHYKVPMSANLKVAK 450

RESULT 5
AA545453
ID AA545453 PRELIMINARY; PRT; 450 AA.
AC AA545453;
```



Best Local Similarity 49.9%; Pred. No. 1.1e-43;  
Matches 231; Conservative 60; Mismatches 102; Indels 70; Gaps 15;

QY 8 NNTSSPGLFQSGDNGLGHNANSALGQPIDRQIEQMAQLLAELLSKLL-----59  
Db 61 NTPSSDTG--SSQOAGI-GNGSSALGQN-----GSAADMLKLEALIPSKNGQEA 110  
QY 60 -SPQSGNAATGAGN---DQTTGVNAGGLNGRKGTTAGTTPQSDSQNMLSEMGNGLDQA 115  
Db 111 GNPLSSGSSGAAGNGSGASPLTSSGAGGVGGAQ-----NPEDLSRSLLDQDSAGSALNNA 165  
QY 116 ITPDGGGGGQIGDNPLLKAMKLIARMDGQSDQFGQPGGTGNNSSAGSTSSGSGSPENDL 175  
Db 166 INPTADGGGQSGNDLKLALLEILGNLMDSKQGEFGQP-----QSGSGQGGGSP-----215  
QY 176 SGKAPSGNSPSGNYSPVSTPSPSTPTSLDPPSPPTKAAGG-----STPVTDHDPD 230  
Db 216 STG-AQAASSGGGGSPAA-----PSAPSSVGGNGGAASAPLTAAPT 257  
QY 231 PV--GSAG-----IGAGNSVAPTSAGANQTVLHDTITVKAGQVFDGKGQTFTAGSELGD 282  
Db 258 GVDGGSAASTASTAGAG-PVSPPTASANTVYNDTIKVGPGSEVFDGKGKTFPTASSKLGD 316  
QY 283 GGOSEKPKLFILEDGASLKNVTMGDDGADGHLGYDAKIDNLHVTNVGDAITVKPNSA 342  
Db 317 GGOAEGQKPLFELAQOATLKNVVGDNAAADVHVRGDAKIDNVHVTNVGDAITVKNSN- 375  
QY 343 GKSHVEITNSSPEHASDKILQNLADTNLSVDNVKAKDFGTFVTRTNGGQGNWDLNLSHI 402  
Db 376 GKPAVEITNSSAQASDKIFQLNADANLTIDNFKAKDFGTFVTRTNGGQGNWDLNLSNI 435  
QY 403 SAEDGKFSFKVSDSEGLNVNTSDISGLDVENHYKVPMSANLKV 445  
Db 436 DAENGKFSFKVSDSEGLNVKGNINLTNNVNNHYKVPDSANLQV 478

RESULT 8  
AAS20352 PRELIMINARY; PRT; 479 AA.  
AC AAS20352;  
DT 02-MAR-2004 (T-EMBLrel. 27, Created)  
DT 02-MAR-2004 (T-EMBLrel. 27, Last sequence update)  
DE HRPW.  
GN HRPW.  
OS Pectobacterium atrosepticum.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Pectobacterium.  
OX NCBI\_TaxID=29471;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=SCRI1039;  
RA Holvea M.C., Bell K.S., Hyman L.J., Avrova A.O., Whisson S.C.,  
RA Birch P.R.J., Both I.K.;  
RT "Rapid demonstration of a role early in disease development for the  
RT type III secretion system of Erwinia carotovora subsp. atroseptica  
RT SCRI1039 using a pooled transposon mutation grid."  
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY96066; AAS20352.1; -. 3463E226CDF0406A CRC64;  
SQ SEQUENCE 479 AA; 46706 MW; 3463E226CDF0406A CRC64;

Query Match 43.5%; Score 1005; DB 2; Length 479;  
Best Local Similarity 49.9%; Pred. No. 1.1e-43;  
Matches 231; Conservative 60; Mismatches 102; Indels 70; Gaps 15;

QY 8 NNTSSPGLFQSGDNGLGHNANSALGQPIDRQIEQMAQLLAELLSKLL-----59  
Db 61 NTPSSDTG--SSQOAGI-GNGSSALGQN-----GSAADMLKLEALIPSKNGQEA 110  
QY 60 -SPQSGNAATGAGN---DQTTGVNAGGLNGRKGTTAGTTPQSDSQNMLSEMGNGLDQA 115  
Db 111 GNPLSSGSSGAAGNGSGASPLTSSGAGGVGGAQ-----NPEDLSRSLLDQDSAGSALNNA 165

QY 116 ITPDGGGGGQIGDNPLLKAMKLIARMDGQSDQFGQPGGTGNNSSAGSTSSGSGSPENDL 175  
Db 166 INPTADGGGQSGNDLKLALLEILGNLMDSKQGEFGQP-----QSGSGQGGGSP-----215  
QY 176 SGKAPSGNSPSGNYSPVSTPSPSTPTSLDPPSPPTKAAGG-----STPVTDHDPD 230  
Db 216 STG-AQAASSGGGGSPAA-----PSAPSSVGGNGGAASAPLTAAPT 257  
QY 231 PV--GSAG-----IGAGNSVAPTSAGANQTVLHDTITVKAGQVFDGKGQTFTAGSELGD 282  
Db 258 GVDGGSAASTASTAGAG-PVSPPTASANTVYNDTIKVGPGSEVFDGKGKTFPTASSKLGD 316  
QY 283 GGOSEKPKLFILEDGASLKNVTMGDDGADGHLGYDAKIDNLHVTNVGDAITVKPNSA 342  
Db 317 GGOAEGQKPLFELAQOATLKNVVGDNAAADVHVRGDAKIDNVHVTNVGDAITVKNSN- 375  
QY 343 GKSHVEITNSSPEHASDKILQNLADTNLSVDNVKAKDFGTFVTRTNGGQGNWDLNLSHI 402  
Db 376 GKPAVEITNSSAQASDKIFQLNADANLTIDNFKAKDFGTFVTRTNGGQGNWDLNLSNI 435  
QY 403 SAEDGKFSFKVSDSEGLNVNTSDISGLDVENHYKVPMSANLKV 445  
Db 436 DAENGKFSFKVSDSEGLNVKGNINLTNNVNNHYKVPDSANLQV 478

RESULT 9  
Q79AW7 PRELIMINARY; PRT; 138 AA.  
AC Q79AW7;  
DT 05-JUL-2004 (T-EMBLrel. 27, Created)  
DT 05-JUL-2004 (T-EMBLrel. 27, Last sequence update)  
DT 05-JUL-2004 (T-EMBLrel. 27, Last annotation update)  
DE HRPW (Fragment).  
GN Name=hrpw;  
OS Erwinia amylovora.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Erwinia.  
OX NCBI\_TaxID=552;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Es321;  
RX MEDLINE=98115919; PubMed=9448330;  
RA Bogdanove A.J., Kim J.F., Wei Z., Kolchinsky P., Charkowski A.O.,  
RA Conlin A.K., Collmer A., Beer S.V.;  
RT "Homology and functional similarity of an hrp-linked pathogenicity  
RT locus, depeF, of Erwinia amylovora and the avirulence locus avrE of  
RT Pseudomonas syringae pathovar tomato."  
RL Proc. Natl. Acad. Sci. U.S.A. 95:1325-1330(1998).  
DR EMBL; U97504; AAC04849.1; -.  
FT NON TER 138  
SQ SEQUENCE 138 AA; 13788 MW; 4FE177177C74B3C6 CRC64;

Query Match 30.5%; Score 705; DB 2; Length 138;  
Best Local Similarity 100.0%; Pred. No. 6e-29;  
Matches 138; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSILTLNNTSSPGLFQSGDNGLGHNANSALGQPIDRQIEQMAQLLAELLSKLLS 60  
Db 1 MSILTLNNTSSPGLFQSGDNGLGHNANSALGQPIDRQIEQMAQLLAELLSKLLS 60  
QY 61 POSGNAATGAGNDQTTGVNAGGLNGRKGTTAGTTPQSDSQNMLSEMGNGLDQAITPDG 120  
Db 61 POSGNAATGAGNDQTTGVNAGGLNGRKGTTAGTTPQSDSQNMLSEMGNGLDQAITPDG 120  
QY 121 QGGGQIGDNPLLKAMLKL 138  
Db 121 QGGGQIGDNPLLKAMLKL 138

RESULT 10  
Q8RP12 PRELIMINARY; PRT; 424 AA.  
ID Q8RP12  
AC Q8RP12;





Db 246 FTADKSGNGDQGENQKPMFELAGATLKNNVLGENEVDGIHVAKNAQEVITINVAQN 305  
Qy 330 VGEDAITVKNPSAGKSHVEITNSPEHASDKILOINADTNLSVDNVRKADPGFTFVRING 389  
Db 306 VGEDLITVKGGAANTNINIKNSAGKADDDKVQVQNLANTHLKIDNFKADDFGTWVRTG 365  
Qy 390 GQO-GNWDNLNLSHSAEDCKFSVKSDEGLNVNTSDISLGDEVNHY 435  
Db 366 GQOFDDMSIELNGIEANHGKFAVKSDSDLLKATGNIAMTVDVHAY 412

RESULT 14  
Q849P3  
ID Q849P3 PRELIMINARY; PRT; 386 AA.  
AC Q849P3;  
DT 01-JUN-2003 (TremBLrel. 24, Created)  
DT 01-JUN-2003 (TremBLrel. 24, Last sequence update)  
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)  
DE Harpin.  
GN Name=hrpW;  
OS Pseudomonas syringae (pv. phaseolicola).  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
OC Pseudomonadaceae; Pseudomonas.  
OX NCBI\_TaxID=319;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Tsaltas D.;  
RT "Biochemical, structural and molecular characterization of resistant  
interactions between Pseudomonas syringae pv. phaseolicola and  
phaseolus vulgaris.";  
RL Thesis (2003), University of London Library, Senate House, Malet  
Street, London, In press.  
DR EMBL; AY210846; AAC50075.1; -.  
DR HSSP; Q9RHW0; 1EE6.  
DR GO; GO:0005576; C:extracellular; IEA.  
DR GO; GO:0030570; F:pectate lyase activity; IEA.  
DR InterPro; IPR011050; Pectin\_lyase\_like.  
DR InterPro; IPR004898; Pect\_lyase.  
DR Pfam; PF03211; Pectate\_lyase; 1.  
SQ SEQUENCE 386 AA; 40459 MW; 1F079531BC895228 CRC64;

Query Match 20.5%; Score 472.5; DB 2; Length 386;  
Best Local Similarity 32.1%; Pred. No. 1.4e-16;  
Matches 141; Conservative 52; Mismatches 121; Indels 125; Gaps 17;  
Qy 16 LFQSGDNLGL-GHNANSALGQPIDROT-----IEQMAQLLAELLSKILLSPOSGNA 66  
Db 47 LFNSAAPKNVSGQPDNSV--QSPQDANNANPESNLKFLSALIVSLQLMLN-LNKQ 103  
Qy 67 ATGAGGNDQTTGVGNAGLNGRKGTTAGTTPOSQNNMLSEMGNNGLDQAITPDGQGGQI 126  
Db 104 DTDQSSSEWQDFQNGGLG-----TFSTDS-----GDSGDGTLAATGDCG--- 145  
Qy 127 GNPILLKAMLLIARMMDGSDQDQPGCTGNSSASCTSSGSGSPNDLSGKAPGSGNSP 186  
Db 146 GDTP-----TATGNN-----TPSVEGSSND--DGVTFQLAMP 176  
Qy 187 SGNYSVSTSPSTPTSPDLPSSPTKAAGSTPTVTHDPDVGSAIGAGNSVAFT 246  
Db 177 N-----HTSGTGPVSD----- 187  
Qy 247 SAGANQ-----TVLHDTITVKAGVFGKGTPTAGSELGDGQSENQKPLFLEDGASL 301  
Db 188 TAGANDQAGKVIIVKDTIKVAANTVYDAGATFTADKSMGNGDQSENQKPLFELAKGATL 247  
Qy 302 KNTVMGDGADGHLVG-DAK-----INLHVTNVGEDAITVKNPSAGKSHVEITNSSPEH 357  
Db 248 KNAHGENEVDGIHVAKNAQEVITINVAENNVGEMITVKEGGAKVTNLIQKNSSAEN 307  
Qy 358 ASDKILQLNADTNLSVDNVRKADPGFTFVRINGGQO-GNWDNLNLSHSAEDCKFSVKS 416  
Db 308 ADDKVFQLNANTHLNITGEANNFGTLVRINGEKQFDMDNKLKIDNFKADDFGTWVRTG 367

Qy 417 EGLNVNTSDISLGDEVNHY 435  
Db 368 E-----DLQLGQQHRH 379

RESULT 15  
Q8KKZ2  
ID Q8KKZ2 PRELIMINARY; PRT; 307 AA.  
AC Q8KKZ2;  
DT 01-OCT-2002 (TremBLrel. 22, Created)  
DT 01-OCT-2002 (TremBLrel. 22, Last sequence update)  
DT 01-JUN-2003 (TremBLrel. 24, Last annotation update)  
DE Hypothetical protein hrpW.  
GN Name=hrpW;  
OS Rhizobium etli.  
OC Plasmid symbiotic plasmid p42d.  
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Rhizobium.  
OX NCBI\_TaxID=29449;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA STRAIN=CFN42;  
RX MEDLINE=91193195; PubMed=2013564;  
RA Girard M.L., Flores M., Brom S., Romero D., Palacios R., Davila G.;  
RT "Structural complexity of the symbiotic plasmid of Rhizobium  
leguminosarum bv. phaseoli.";  
RL J. Bacteriol. 173:2411-2419 (1991).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA STRAIN=CFN42;  
RX MEDLINE=97419521; PubMed=9274036;  
RA Ramirez-Romero M.A., Bustos P., Girard L., Rodriguez O.,  
RA Cevallos M.A., Davila G.;  
RT "Sequence, localization and characteristics of the replicator region  
of the symbiotic plasmid of Rhizobium etli.";  
RL Microbiology 143:2825-2831 (1997).  
RN [3]  
RP SEQUENCE FROM N.A.  
RA STRAIN=CFN42;  
RX MEDLINE=22309397; PubMed=12421308;  
RA Quintero V., Cevallos M.A., Davila G.;  
RT "A site-specific recombinase (RinQ) is required to exert  
incompatibility towards the symbiotic plasmid of Rhizobium etli.";  
RL Mol. Microbiol. 46:1023-1032 (2002).  
RN [4]  
RP SEQUENCE FROM N.A.  
RA STRAIN=CFN42;  
RX Ramirez M.A., Bustos P., Girard L., Rodriguez O., Cevallos M.A.,  
RA Davila G.;  
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RP SEQUENCE FROM N.A.  
RA STRAIN=CFN42;  
RX Quintero V., Bustos P., Davila G.;  
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.  
RN [6]  
RP SEQUENCE FROM N.A.  
RA STRAIN=CFN42;  
RX Gonzalez V., Bustos P., Medrano-Soto A., Ramirez-Romero M.A.,  
RA Romero D., Salgado H., Hernandez-Gonzalez I., Hernandez-Celis J.C.,  
RA Quintero V., Girard L., Rodriguez O., Flores M., Cevallos M.A.,  
RA Collado-Vides J., Davilla G.;  
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
RN [7]  
RP SEQUENCE FROM N.A.  
RA STRAIN=CFN42;  
RX Quintero V., Bustos P., Davila G.;  
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U06928; U06928.1; -.  
DR HSSP; Q9RHW0; 1EE6.  
DR GO; GO:0005576; C:extracellular; IEA.  
DR GO; GO:0030570; F:pectate lyase activity; IEA.  
DR InterPro; IPR011050; Pectin\_lyase\_like.

```
DR InterPro: IPR004898; Pect. lyase.
DR Pfam: PF03211; Pectate lyase; 1.
KW Hypothetical protein; Plasmid.
SQ SEQUENCE 307 AA; 32912 MW; ED8844FA3AJA5E98 CRC64;

Query Match 18.3%; Score 423.5; DB 2; Length 307;
Best Local Similarity 37.7%; Pred. No. 3.6e-14;
Matches 121; Conservative 48; Mismatches 83; Indels 69; Gaps 14;

QY 144 DGSDQGPQGTGNSSASGTS---SSGSGPFNDLSCGKAPSGNSPGNSPVSTSPPS 200
DB :||:|||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 45
DB DLESDCLG---GSNRADTFTTIRTKVSPFEDF-----SGNPPT-----LTSYVPS 66
QY 201 TPTS-----PSPDLF-----PSPFKAAGSTPTVDPDPVGSAGIAGN 241
DB :||:|||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 114
DB RETSENGMDSPDKLLRKHHINWQSDSKVDPSEKQATTLQT-TTEKPD----- 114
QY 242 SVAPTSAGANTVLHDTITVKAGQVFDGKQGTFTAGSELGDDGQSENOKPLFILEDCASL 301
DB :||:|||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 169
DB 115 -----LSKEGSIIVNEPIVVDGG-VFDGKATYFASKLGGQSGSETQSPFLFKNGAIL 169
QY 302 KNYTMGDGAGIHLGDAKIDNLHVTNVGDAITVKPNSAGKSKSHVEITNSSFEHASDK 361
DB :||:|||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 224
DB 170 KNVDLGNGAGIHHVYDGTALFNWQNVGEDALTVK---SAG---DITLIIGSAKGATDK 224
QY 362 ILQINADTNLSVNVKAKDFTFTVNTGGQGGWDLNLSHISABDGKFS-----FVKSDSE 417
DB :||:|||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 279
DB 225 IFQINADTFYKDFVADGFTTLVRTNGKQIDAD-----VTIDGAFSHGSNVFRDSS 279
QY 418 GLNVN-TSDISLGDVENHYKV 437
DB :||:|||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 280
DB LASVTFSLSDITLDDVKNWTRV 300

RESULT 16
Q8XVQ5 PRELIMINARY; PRT; 380 AA.
AC Q8XVQ5
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE PROBABLE HARPIN-RELATED PROTEIN.
GN Names: popW; Synonyms: R500072; Ordered locus names: R5C2775;
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Ralstonia.
OX NCBI_TaxID=305;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GMI1000;
RX MEDLINE=21691879; PubMed=11823852;
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA Arlat M., Billault A., Brottier P., Camus J.C., Cattolico L.,
RA Chandler M., Choise N., Claudel-Renard C., Cunnac S., Demange N.,
RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schlex T.,
RA Siguer P., Thebaud P., Whalen M., Wincker P., Levy M.,
RA Weissenbach J., Boucher C.A.;
RT "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RL Nature 415:497-502(2002)
DR ENBL; AL646071; CAD16482.1; -.
DR HSBP; Q9RHW0; IEE6.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0030570; F:pectate lyase activity; IEA.
DR InterPro; IPR011050; F:pectate lyase activity; IEA.
DR InterPro; IPR004898; Pect. lyase.
DR Pfam; PF03211; Pectate lyase; 1.
KW Complete proteome.
SQ SEQUENCE 380 AA; 39809 MW; 06777197E97365D1 CRC64;

Query Match 18.0%; Score 415.5; DB 2; Length 380;
Best Local Similarity 30.5%; Pred. No. 1.2e-13;
Matches 143; Conservative 52; Mismatches 133; Indels 141; Gaps 20;

QY 8 NNTSSSPGLFQSGDNGLGGHNANSALGQOPIDQRTIEQMAQLLAELKLLSPQSGNAA 67
DB :||:|||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 45
DB 9 NNHFQTPSTW-----NHDAGSS-----IDTSLQRAVQLDDQVLQQL----- 45
QY 68 TGAGGNDQTTGVGNAGLNGRKGTAGTTPOSQSNMLSEMG---NNGLDQAITPDGQGGGQ 125
DB :||:|||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 75
DB 46 -----EARKLFG-----NMLNPGADNAGQNH-----GGHGGGH 75
QY 126 IGDNPLLKAMLUKLIARMDGQSQDFQPGPT-GNNSASSGTSSSGSGSPFNDLSGK-APSG 183
DB :||:|||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 117
DB 76 -----HGGNGFGNGRFGSPHANSAPQDLELPANKPNNGKHNTSA 117
QY 184 NSPSGNSPVSTPSPSTPTSPSPLDFPSSPTKAAGGSTPTVDHDPVGSAGIAGNSV 243
DB :||:|||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 165
DB 118 STPTDQTAPST--SPTTGTSP--PSTTSATEGKAYGVKVP-----PEPTGVVDVS----- 165
QY 244 AFTSAGANTVLHDTITVKAGQVFDGKQGTFTAGSELGDDGQSENOKPLFILEDCASLKN 303
DB :||:|||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 212
DB 166 -----KPIVVRKAGETFDGGKYRPTKEMGDSQNEHQKPLFIPSGATLKN 212
QY 304 VTMGDGAGIHLGDAKIDNLHVTNVGDAITVKPNSAGKSKH----- 347
DB :||:|||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 268
DB 213 VQY--SGDGIHLLGSAKLDRVNVNQVGDAIT--DGAKNRAHDAKIAIGIDPASIPGGT 268
QY 348 --VEITNSSFEHASDKILQINADTNLSVNVKAKDFTFTVNTGG-----CQGNW 395
DB :||:|||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 328
DB 269 PRVEITNSAPYGAOKLAQINGDVIDQVKGMYNAGKVFRTNGGDTQIKATVNVQDSNF 328
QY 396 DLNLSHISABDGKFS-----VKSDS--EGLNVNTSDISLGDVENHYK 436
DB :||:|||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 376
DB 329 QNVSEAVFRDTSKFTASPSDDVKSADPFGLADPKSQVT-GTNKVSYK 376

RESULT 17
Q9RHW0 PRELIMINARY; PRT; 224 AA.
AC Q9RHW0
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Pectate lyase.
OS Bacillus sp.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1409;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KSM-P15;
RX MEDLINE=2023260; PubMed=10759850;
RA Hatada Y., Saito K., Koike K., Yoshimatsu T., Ozawa T., Kobayashi T.,
RA Ito S.;
RT "Deduced amino-acid sequence and possible catalytic residues of a
RT novel pectate lyase from an alkaliphilic strain of Bacillus.";
RL Eur. J. Biochem. 267:2268-2275(2000).
DR ENBL; AB011839; BAA87892.1; -.
DR PDB; 1ER6; X-ray; A=28-224;
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0016829; F:lyase activity; IEA.
DR GO; GO:0030570; F:pectate lyase activity; IEA.
DR InterPro; IPR011050; Pectin lyase like.
DR InterPro; IPR004898; Pect. lyase.
DR Pfam; PF03211; Pectate lyase; 1.
KW Lyase.
SQ SEQUENCE 224 AA; 23751 MW; 99D04821B09DE523 CRC64;

Query Match 15.7%; Score 362.5; DB 2; Length 224;
Best Local Similarity 46.1%; Pred. No. 3.3e-11;
Matches 83; Conservative 29; Mismatches 59; Indels 9; Gaps 4;

QY 248 AGANTVLHDTITVKAGQVFDGKQGTFTAG-SELGDDGQSENOKPLFILEDCASLKNVTV 306
DB :||:|||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 84
DB 25 AEAAPTVVHETIRVPAGQTFDGKQTVVAVNPNTLGDGSAQENQKPIFRLGASLKNVVI 84
QY 307 GDDGAGIHLGDAKIDNLHVTNVGDAITVKPNSAGKSKSHVEITNSSFEHASDKILQLN 366
```

```

Db      85  GAPAADGVHCYGDCTTINVIWEDVGEDALTLK--SSGT-----VNISGAAYKAYDKVQFIN 139
QY      367  ADTLNSVDNVKAKDGFVTRTNGGQGNWDLNLSHISAEDGKFSFKSDSE---GLNVT 423
Db      140  AAGTINIRNPADIDIGKLVQNGGTTVKVMVENCNISRVKDALIRLDSSTSGRIYNT 199

RESULT 18
OBPHIS
ID      08PHIS PRELIMINARY; PRT; 303 AA.
AC      08PHIS;
DT      01-OCT-2002 (T-EMBLrel. 22, Created)
DT      01-OCT-2002 (T-EMBLrel. 22, Last sequence update)
DE      HrpW protein.
GN      Name=hrpW; OrderedLocusNames=XAC2922;
OS      Xanthomonas axonopodis (pv. citri).
OC      Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OX      Xanthomonadaceae; Xanthomonas.
NCBI_TaxID=92829;
[1]
SEQUENCE FROM N.A.
RX      STRAIN=306 / ATCC 13902 / XV 101;
RX      MEDLINE=2202245; PubMed=12024217;
RX      da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RX      Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RX      Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RX      Camarotte G., Cannavan P., Cardoso J., Chambergo F., Ciapina L.P.,
RX      Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorri H.,
RX      Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RX      Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RX      Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RX      Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RX      Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RX      Pereira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RX      Spinola L.A.P., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RX      Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RX      Setubal J.C., Kitajima J.P.;
RX      "Comparison of the genomes of two Xanthomonas pathogens with differing
RX      host specificities.";
RX      Nature 417:459-463 (2002).
RX      EMBL; AB011934; AAC37767.1; --
RX      HSP; Q9RHW0; 1EE6.
RX      GO; GO:0005576; C:extracellular; IEA.
RX      GO; GO:0030570; F:pectate lyase activity; IEA.
RX      InterPro; IPR011050; Pectin lyase like.
RX      Pfam; PF03211; Pectate lyase.
RX      Complete proteome.
SQ      SEQUENCE 303 AA; 31949 MW; 76E283B23852001B CRC64;

Query Match.
Best Local Similarity 30.7%; Pred. No. 1.2e-10;
Matches 100; Conservative 45; Mismatches 85; Indels 96; Gaps 14;

QY      135  MKLIARMDQSQDFQCPGTG-----NNSASGTSSSGGSPNDLSGGKAPGSPGNY 190
Db      3  MSELLQRLI---QTQFQFQNTSQVDWNPSPSHCTSGNQ-----EGNEHGR- 46
QY      191  SPVSTFPSPPTSPSTPLDFPSPKAG-----STPVT-DHPDP-----VGSAGIG 238
Db      47  -----RNGRGDNFLATTPTSDSHSDKSAHSTVTGSPVG 81
QY      239  AGNSVAFTSAGANQTVLHDTITVKAGQVFDGKGQFTTAGSELGDGQSEKQKPIFILEDG 298
Db      82  SGRITNVNS-----PIVVHKEGVDFGHNNLYVGGSGIGDGSQSEHQPFVVEQG 131
QY      299  ASLKNVTMGDDGADGHLHYGDAKIDNLHVTNVGDAITVKNPSAGKKSH----- 347
Db      132  GTLQVIRM--SGGDGHLHLLGDLATLKNVNLNVSEDAWTI--DGFGNREHDSRIAGTPASG 198

```

```

QY      348  -----VEITNSFPHASDKILQNLNADTNLSVDNVKAKDGFVTRTNGGQGNWDLNLSH 401
Db      189  LPAPKPIELDSFQNASDKVIQDQADVLLRNVSNGAKVKFRTNG---GHTDID-SH 244
QY      402  ISAEQDKFSFKV-----SDSEGLNVN 422
Db      245  VIVEDSTLKGKAEAVFRTDAPGAHVS 270

RESULT 19
O34310
ID      O34310 PRELIMINARY; PRT; 221 AA.
AC      O34310; Q795F2;
DT      01-JAN-1998 (T-EMBLrel. 05, Created)
DT      01-JAN-1998 (T-EMBLrel. 05, Last sequence update)
DT      01-OCT-2004 (T-EMBLrel. 28, Last annotation update)
DE      YvpA.
GN      Name=yvpA; OrderedLocusNames=BSU34950;
OS      Bacillus subtilis.
OC      Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX      NCBI_TaxID=1423;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=98230327; PubMed=9570401;
RX      Reizer J., Hoischen C., Tilgemyer F., Rivolta C., Rabus R.,
RX      Stulke J., Karamata D., Sailer M.H. Jr., Hillen W.,
RX      "A novel protein kinase that controls carbon catabolite repression in
RX      bacteria.";
RX      Mol. Microbiol. 27:1157-1169 (1998).
[2]
SEQUENCE FROM N.A.
RX      Lazarevic V., Soldo B., Rivolta C., Reynolds S., Maue C.,
RX      Karamata D.;
RX      Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
[3]
SEQUENCE FROM N.A.
RX      STRAIN=168;
RX      MEDLINE=98044033; PubMed=9384377; DOI=10.1038/36786;
RX      Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RX      Azevedo V., Bertero M., Bessieres P., Bolotin A., Borchert S.,
RX      Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RX      Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RX      Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RX      Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
RX      Entian K.-D., Errington J., Fabret C., Ferracioli S., Foulger D.,
RX      Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RX      Ghim S.Y., Glaeser P., Goffeau A., Gollightly E.J., Grandi G.,
RX      Guilserri G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
RX      Hilbert H., Holsappel S., Hosono S., Hulio M.F., Itaya M.,
RX      Jones L.-M., Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M.,
RX      Klein C., Kobayashi Y., Koetter P., Koningsstein G., Krohn S.,
RX      Kumano M., Kurita K., Lapidus A., Lardinois S., Lauber J.,
RX      Lazarevic V., Lee S.M., Levine A., Liu H., Masuda S., Mauel C.,
RX      Madigou C., Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S.,
RX      Noback M., Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B.,
RX      Park S.H., Parro V., Pohl T.M., Portetelle D., Porwollik S.,
RX      Prescott A.M., Prescan E., Pujic P., Purnelle B., Rapoport G.,
RX      Rose M., Sadate Y., Sato T., Scanlan E., Schleich S., Schroeter R.,
RX      Scoffone F., Sekiguchi J., Sekowska A., Seror S.J., Seror P.,
RX      Shin B.S., Soldo B., Sorokin A., Tacconi E., Takagi T., Takahashi H.,
RX      Takemaru K., Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P.,
RX      Tognoni A., Tosato V., Uchiyama S., Vandenbol M., Vannier F.,
RX      Vassarotti A., Viari A., Wambut R., Wedler E., Wedler H.,
RX      Weitzenecker T., Winters P., Wipat A., Yamamoto H., Yamane K.,
RX      Yasumoto K., Yata K., Yoshida K., Yoshikawa H.F., Zumstein E.,
RX      Yoshikawa H., Zanchin A.;
RX      "The complete genome sequence of the Gram-positive bacterium Bacillus
RX      subtilis.";
RX      Nature 390:249-256 (1997).
RL      EMBL; AF017113; AAC67291.1; --
DR      EMBL; Z99121; CAB15500.1; --
DR      PIR; A70045; A70045.

```







GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2005, 12:04:36 ; Search time 20 Seconds  
(without alignments)  
1482.207 Million cell updates/sec

Title: US-09-879-248-6

Perfect score: 2310

Sequence: 1 MSILTLNNTSSPGLFQSG.....LGDVENHYKVPMSNLKVAE 447

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 56318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : Issued Patents AA:\*

- 1: /cgn2\_6/prodata/1/iaa/5A\_COMB.pep.\*
- 2: /cgn2\_6/prodata/1/iaa/5B\_COMB.pep.\*
- 3: /cgn2\_6/prodata/1/iaa/6A\_COMB.pep.\*
- 4: /cgn2\_6/prodata/1/iaa/6B\_COMB.pep.\*
- 5: /cgn2\_6/prodata/1/iaa/PCUTUS\_COMB.pep.\*
- 6: /cgn2\_6/prodata/1/iaa/backfilese1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2310	100.0	447	3	US-09-120-927-2
2	2310	100.0	447	4	US-09-431-614-6
3	559	24.2	424	3	US-09-120-817-2
4	559	24.2	424	4	US-09-431-614-14
5	359.5	15.6	197	3	US-09-402-668-2
6	341	14.8	221	3	US-09-198-956-4
7	341	14.8	221	4	US-09-670-141-4
8	190.5	8.2	62	3	US-09-402-668-10
9	181.5	7.9	2411	3	US-09-268-347-36
10	174	7.5	2042	4	US-09-077-098A-6
11	165.5	7.2	344	1	US-08-891-254-7
12	165.5	7.2	344	2	US-08-819-539-7
13	165.5	7.2	344	2	US-09-030-270A-7
14	165.5	7.2	344	3	US-08-984-207-7
15	165.5	7.2	344	3	US-09-013-587-7
16	165.5	7.2	344	4	US-09-086-118-27
17	165.5	7.2	344	4	US-09-431-614-15
18	165.5	7.2	344	5	PCT-US96-08819-7
19	160.5	6.9	907	2	US-09-010-928B-4
20	157.5	6.8	2870	4	US-09-479-467A-15
21	157.5	6.8	3178	4	US-09-479-467A-4
22	156.5	6.8	1912	1	US-08-409-995-4
23	156.5	6.8	1912	3	US-08-685-467-4
24	156	6.8	2039	4	US-09-077-098A-7
25	155.5	6.7	2353	3	US-09-377-155-33
26	155.5	6.7	2353	4	US-08-313-942-4
27	155.5	6.7	2353	3	US-09-669-974-33
28	155.5	6.7	2353	4	US-09-797-862-33
29	155.5	6.7	2353	4	US-09-684-707-4
30	155.5	6.7	2354	3	US-09-268-347-47
31	155	6.7	385	5	PCT-US93-06243-2
32	155	6.7	495	2	US-08-794-795-2
33	155	6.7	495	3	US-09-249-200-2
34	154.5	6.7	975	4	US-09-328-352-4764
35	154	6.7	571	3	US-09-134-001C-3865
36	151	6.5	385	1	US-08-891-254-3
37	151	6.5	385	2	US-08-819-539-3
38	151	6.5	385	5	PCT-US96-08819-3
39	151	6.5	403	2	US-08-200-724A-2
40	151	6.5	403	2	US-09-030-270A-3
41	151	6.5	403	3	US-08-851-376A-2
42	151	6.5	403	3	US-08-984-207-3
43	151	6.5	403	3	US-09-013-587-3
44	151	6.5	403	4	US-09-086-118-23
45	151	6.5	403	4	US-09-431-614-3
46	151	6.5	619	4	US-09-252-991A-26352
47	151	6.5	889	3	US-09-336-447A-15
48	151	6.5	889	4	US-09-952-267B-15
49	150.5	6.5	745	2	US-09-010-928B-28
50	150.5	6.5	870	2	US-09-010-928B-2
51	150	6.5	448	4	US-09-248-796A-19135
52	150	6.5	520	2	US-08-794-795-6
53	150	6.5	520	3	US-09-249-200-6
54	147.5	6.4	489	2	US-08-794-795-7
55	147.5	6.4	489	3	US-09-249-200-7
56	147.5	6.4	518	1	US-08-392-367B-2
57	147.5	6.4	518	3	US-08-893-467A-2
58	146.5	6.3	943	3	US-09-056-556-204
59	146.5	6.3	943	4	US-09-072-596-199
60	146.5	6.3	943	4	US-09-477-135A-131
61	146.5	6.3	943	4	US-09-072-967-204
62	146.5	6.3	1180	4	US-09-206-942-65
63	146.5	6.3	1188	4	US-09-206-942-63
64	146.5	6.3	1776	4	US-09-556-877-179
65	146.5	6.3	1776	4	US-09-620-412C-179
66	146.5	6.3	1776	4	US-09-598-419-179
67	145	6.3	334	3	US-09-060-756-728
68	145	6.3	334	4	US-09-670-314-728
69	144.5	6.2	2736	4	US-09-252-991A-30227
70	144	6.2	1581	3	US-09-110-517-2
71	143.5	6.2	679	3	US-08-913-942-15
72	143.5	6.2	679	3	US-09-268-347-26
73	143.5	6.2	679	4	US-09-684-707-15
74	142.5	6.2	1004	3	US-09-268-347-30
75	141.5	6.1	541	4	US-09-538-092-347
76	141.5	6.1	2315	4	US-09-543-681A-5434
77	140.5	6.1	412	4	US-09-248-796A-26345
78	140.5	6.1	1060	3	US-08-911-393-2
79	140.5	6.1	1060	4	US-09-955-909-2
80	140.5	6.1	1146	4	US-09-198-452A-580
81	138.5	6.0	1036	4	US-09-206-942-73
82	138.5	6.0	1475	4	US-09-538-092-1160
83	138.5	6.0	1477	4	US-09-206-942-71
84	138	6.0	353	2	US-08-687-702-37
85	138	6.0	2090	4	US-09-538-092-1081
86	137.5	6.0	231	4	US-09-248-796A-14281
87	137.5	6.0	481	4	US-09-724-623-77
88	137	5.9	300	3	US-08-765-856-2
89	137	5.9	300	3	US-08-935-009A-2
90	137	5.9	842	5	PCT-US96-02331-15
91	137	5.9	1268	4	US-09-270-767-45320
92	136.5	5.9	681	4	US-09-490-291-4
93	136.5	5.9	691	4	US-09-490-291-6
94	136.5	5.9	1477	1	US-08-038-682-4
95	136.5	5.9	1477	1	US-08-302-832-4
96	136.5	5.9	1477	2	US-08-530-198-4
97	136.5	5.9	1477	2	US-08-469-880-4
98	136.5	5.9	1477	2	US-08-728-470-4
99	136.5	5.9	1477	2	US-08-617-697-4
100	136.5	5.9	1477	3	US-08-719-641-4

Sequence 33, Appli  
Sequence 4, Appli  
Sequence 47, Appli  
Sequence 2, Appli  
Sequence 2, Appli  
Sequence 2, Appli  
Sequence 4764, Ap  
Sequence 3865, Ap  
Sequence 3, Appli  
Sequence 3, Appli  
Sequence 3, Appli  
Sequence 3, Appli  
Sequence 2, Appli  
Sequence 3, Appli  
Sequence 23, Appli  
Sequence 3, Appli  
Sequence 2632, A  
Sequence 15, Appli  
Sequence 15, Appli  
Sequence 28, Appli  
Sequence 2, Appli  
Sequence 19135, A  
Sequence 6, Appli  
Sequence 7, Appli  
Sequence 2, Appli  
Sequence 2, Appli  
Sequence 204, App  
Sequence 199, App  
Sequence 131, App  
Sequence 204, App  
Sequence 65, Appli  
Sequence 63, Appli  
Sequence 179, App  
Sequence 179, App  
Sequence 179, App  
Sequence 728, App  
Sequence 728, App  
Sequence 30227, A  
Sequence 2, Appli  
Sequence 15, Appli  
Sequence 15, Appli  
Sequence 26, Appli  
Sequence 30, Appli  
Sequence 347, App  
Sequence 5434, Ap  
Sequence 26345, A  
Sequence 2, Appli  
Sequence 2, Appli  
Sequence 580, App  
Sequence 73, Appli  
Sequence 1160, Ap  
Sequence 71, Appli  
Sequence 37, Appli  
Sequence 1081, Ap  
Sequence 14281, A  
Sequence 77, Appli  
Sequence 2, Appli  
Sequence 2, Appli  
Sequence 15, Appli  
Sequence 45320, A  
Sequence 4, Appli  
Sequence 6, Appli  
Sequence 4, Appli  
Sequence 4, Appli  
Sequence 4, Appli  
Sequence 4, Appli  
Sequence 4, Appli  
Sequence 4, Appli

ALIGNMENTS

RESULT 1

US-09-120-927-2  
; Sequence 2, Application US/09120927  
; Patent No. 6262018  
; GENERAL INFORMATION:  
; APPLICANT: Kim, Jihyun Francis  
; APPLICANT: Beer, Steven V.  
; TITLE OF INVENTION: HYPERSENSITIVE RESPONSE ELICITOR FROM  
; TITLE OF INVENTION: ERWINIA AMYLOVORA AND ITS USE  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Nixon, Hartgrave, Devans & Doyle LLP  
; STREET: P.O. Box 1051, Clinton Square  
; CITY: Rochester  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 14603  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/120,927  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/055,108  
; FILING DATE: 06-AUG-1977  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Goldman, Michael L.  
; REGISTRATION NUMBER: 30,727  
; REFERENCE/DOCKET NUMBER: 19603/1581  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (716) 263-1304  
; TELEFAX: (716) 263-1600  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 447 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-120-927-2

Query Match 100.0%; Score 2310; DB 3; Length 447;  
Best Local Similarity 100.0%; Pred. No. 5.8e-175;  
Matches 447; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MSILTLNNNTSSSPGLFQSGDNGLGHNANSALGQOPIDRTIQMAQLLAELLS 60  
DB 1 MSILTLNNNTSSSPGLFQSGDNGLGHNANSALGQOPIDRTIQMAQLLAELLS 60  
QY 61 POSGNAATGAGGNDQTTGVGNAGGLNGRKTAGTTPOSQSNMLSEMGNGLDQAITPDG 120  
DB 61 POSGNAATGAGGNDQTTGVGNAGGLNGRKTAGTTPOSQSNMLSEMGNGLDQAITPDG 120  
QY 121 QGGGQIGDNPFLKAMLLIARMMDGSDQFGPGTGNNSASSGTSSTSSGSPFNDLSGGKA 180  
DB 121 QGGGQIGDNPFLKAMLLIARMMDGSDQFGPGTGNNSASSGTSSTSSGSPFNDLSGGKA 180  
QY 181 PSGNSPSGNSYSPVSTFSPSTPTSPPLDPPSSPTKAAGGSTPVTHDPDVGAGIGAG 240  
DB 181 PSGNSPSGNSYSPVSTFSPSTPTSPPLDPPSSPTKAAGGSTPVTHDPDVGAGIGAG 240  
QY 241 NSVAFTSAGANQTVLHDTITVKAGQVFDGKGTFTAGSELGQSGSENQKPLFILEDGA 300  
DB 241 NSVAFTSAGANQTVLHDTITVKAGQVFDGKGTFTAGSELGQSGSENQKPLFILEDGA 300

QY 301 LKNVTMGDDGADGIHLGYGDAKIDNLHVTNVGDAITVKNSAGKSHVEITNSFEHSD 360  
DB 301 LKNVTMGDDGADGIHLGYGDAKIDNLHVTNVGDAITVKNSAGKSHVEITNSFEHSD 360  
QY 361 KILQLNADTNLSVDNVKAKDFGTFTVNTNGQQQGNWDLNLSHISAEDCKFSFVKSDSEGLN 420  
DB 361 KILQLNADTNLSVDNVKAKDFGTFTVNTNGQQQGNWDLNLSHISAEDCKFSFVKSDSEGLN 420  
QY 421 VNTSDISLGDVENHYKVPMSANLKVAE 447  
DB 421 VNTSDISLGDVENHYKVPMSANLKVAE 447  
RESULT 2  
US-09-431-614-6  
; Sequence 6, Application US/09431614  
; Patent No. 6624139  
; GENERAL INFORMATION:  
; APPLICANT: Wei, Zhong-Min  
; APPLICANT: Schading, Richard L.  
; TITLE OF INVENTION: HYPERSENSITIVE RESPONSE ELICITOR-INDUCED STRESS  
; TITLE OF INVENTION: RESISTANCE  
; FILE REFERENCE: 21829/41 (EBC-003)  
; CURRENT APPLICATION NUMBER: US/09/431,614  
; EARLIER FILING DATE: 1999-11-02  
; EARLIER APPLICATION NUMBER: 60/107,243  
; EARLIER FILING DATE: 1998-11-05  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 6  
; LENGTH: 447  
; TYPE: PRT  
; ORGANISM: Erwinia amylovora  
US-09-431-614-6

Query Match 100.0%; Score 2310; DB 4; Length 447;  
Best Local Similarity 100.0%; Pred. No. 5.8e-175;  
Matches 447; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MSILTLNNNTSSSPGLFQSGDNGLGHNANSALGQOPIDRTIQMAQLLAELLS 60  
DB 1 MSILTLNNNTSSSPGLFQSGDNGLGHNANSALGQOPIDRTIQMAQLLAELLS 60  
QY 61 POSGNAATGAGGNDQTTGVGNAGGLNGRKTAGTTPOSQSNMLSEMGNGLDQAITPDG 120  
DB 61 POSGNAATGAGGNDQTTGVGNAGGLNGRKTAGTTPOSQSNMLSEMGNGLDQAITPDG 120  
QY 121 QGGGQIGDNPFLKAMLLIARMMDGSDQFGPGTGNNSASSGTSSTSSGSPFNDLSGGKA 180  
DB 121 QGGGQIGDNPFLKAMLLIARMMDGSDQFGPGTGNNSASSGTSSTSSGSPFNDLSGGKA 180  
QY 181 PSGNSPSGNSYSPVSTFSPSTPTSPPLDPPSSPTKAAGGSTPVTHDPDVGAGIGAG 240  
DB 181 PSGNSPSGNSYSPVSTFSPSTPTSPPLDPPSSPTKAAGGSTPVTHDPDVGAGIGAG 240  
QY 241 NSVAFTSAGANQTVLHDTITVKAGQVFDGKGTFTAGSELGQSGSENQKPLFILEDGA 300  
DB 241 NSVAFTSAGANQTVLHDTITVKAGQVFDGKGTFTAGSELGQSGSENQKPLFILEDGA 300  
QY 301 LKNVTMGDDGADGIHLGYGDAKIDNLHVTNVGDAITVKNSAGKSHVEITNSFEHSD 360  
DB 301 LKNVTMGDDGADGIHLGYGDAKIDNLHVTNVGDAITVKNSAGKSHVEITNSFEHSD 360  
QY 361 KILQLNADTNLSVDNVKAKDFGTFTVNTNGQQQGNWDLNLSHISAEDCKFSFVKSDSEGLN 420  
DB 361 KILQLNADTNLSVDNVKAKDFGTFTVNTNGQQQGNWDLNLSHISAEDCKFSFVKSDSEGLN 420  
QY 421 VNTSDISLGDVENHYKVPMSANLKVAE 447  
DB 421 VNTSDISLGDVENHYKVPMSANLKVAE 447

only  
2 domains  
claim #1  
3 domains

but  
compensating in  
large wobble  
class  
cl. parent  
(possible 103)  
# wobble also discloses  
active fragment

## RESULT 3

US-09-120-817-2  
; Sequence 2, Application US/09120817  
; Patent No. 6172184  
; GENERAL INFORMATION:  
; APPLICANT: Collmer, Alan  
; APPLICANT: Charkowski, Amy  
; APPLICANT: Alfano, James R.  
; TITLE OF INVENTION: HYPERSENSITIVE RESPONSE ELICITOR FROM  
; TITLE OF INVENTION: PSEUDOMONAS SYRINGAE AND ITS USE  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP  
; STREET: P.O. Box 1051, Clinton Square  
; CITY: Rochester  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 14603  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/120,817  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/055,107  
; FILING DATE: 06-AUG-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Goldman, Michael L.  
; REGISTRATION NUMBER: 30,727  
; REFERENCE/DOCKET NUMBER: 19603/1741  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (716) 263-1304  
; TELEFAX: (716) 263-1600  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 424 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-09-120-817-2

Query Match 24.2%; Score 559; DB 3; Length 424;  
Best Local Similarity 36.1%; Pred. No. 2.9e-36;  
Matches 147; Conservative 54; Mismatches 132; Indels 74; Gaps 13;  
  
QY 37 QPIDRTIEQMAQLLAELKSL---LSPOSQNAATGAGNDQTTGVGNAGLNGRKGATG 93  
Db 72 KPNDQS--NIAKLISALIMSLQLMLTNSKKQDTNQEQPDSQAPQNNGLG-----122  
  
QY 94 TTPQSDSQNMLSEMGNNGLDQAITPGCGGGQIGDNPLLKAMKLIARWMDGSDQFGQP 153  
Db 123 -TPSADS-----GGGG-----TPDATGGG-GDTP-----SATGGG 151  
  
QY 154 GTGNNSASGTSSTSSGGSPFNDLSGGKAPSGNSPGNSYSPVSTPSPSTPTSPSPDPS 213  
Db 152 GGTPTATGGGGGGGTPTATGGG---SGTPTATGGGGGVTPOITPOL-----A 200  
  
QY 214 SPTKAAGGSTPVTDHPDPVGSAGIGAGNSVAFPSAGANQTVLHDTITVKAGQVFDKGQT 273  
Db 201 NPRTSG-----TGSVSDTAGS---TEQAGKINNVKDTIKVGAGEVFDGHGAT 245  
  
QY 274 FTAGSELGCGQSENQKPLFILEDGASLKNVTMGDDGADGIHLG-----DAKIDNLHVTN 329  
Db 246 FTADKSMGNDQGENQKPMFELAEATLKNVNLGENEVDDGIHVKAQNAQEVITDINVAQN 305  
  
QY 330 VGEDAITVKNPSAGKSHVEITNSSPEHASDKILQNLADTNLSVDNVKAKDFCTFVRTNG 389  
Db 306 VGEDLITVKEGGAATVNLNKNSSAKGADDKVQLNANTHLKIDNFKADDFGTWRTNG 365  
  
QY 390 GQO-GNWDNLNLSHISAEDCKFSFKVSDSEGLNVNTSDISLGDVYENHY 435  
Db 366 GKQFDDMSIELNGIEANHGKFAVKSDDSLKLATGNIAMTDVKHAY 412

## RESULT 5

US-09-402-668-2  
; Sequence 2, Application US/09402668  
; Patent No. 6172030  
; GENERAL INFORMATION:  
; APPLICANT: WADA, Yasunao  
; APPLICANT: KASAI, Miyuki  
; APPLICANT: SHIKATA, Shitsuw  
; APPLICANT: SUZUMATSU, Atsushi  
; APPLICANT: KOIKE, Kenzo  
; APPLICANT: HATADA, Yuji  
; APPLICANT: KOBAYASHI, Tohru

QY 390 GQO-GNWDNLNLSHISAEDCKFSFKVSDSEGLNVNTSDISLGDVYENHY 435  
Db 366 GKQFDDMSIELNGIEANHGKFAVKSDDSLKLATGNIAMTDVKHAY 412  
  
RESULT 4  
US-09-431-614-14  
; Sequence 14, Application US/09431614  
; Patent No. 6624139  
; GENERAL INFORMATION:  
; APPLICANT: Wei, Zhong-Min  
; APPLICANT: Schading, Richard L.  
; TITLE OF INVENTION: HYPERSENSITIVE RESPONSE ELICITOR-INDUCED STRESS  
; TITLE OF INVENTION: RESISTANCE  
; FILE REFERENCE: 21829/41 (SBC-003)  
; CURRENT APPLICATION NUMBER: US/09/431,614  
; CURRENT FILING DATE: 1999-11-02  
; EARLIER APPLICATION NUMBER: 60/107,243  
; EARLIER FILING DATE: 1998-11-05  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 14  
; LENGTH: 424  
; TYPE: PRT  
; ORGANISM: Pseudomonas syringae  
; US-09-431-614-14

Query Match 24.2%; Score 559; DB 4; Length 424;  
Best Local Similarity 36.1%; Pred. No. 2.9e-36;  
Matches 147; Conservative 54; Mismatches 132; Indels 74; Gaps 13;  
  
QY 37 QPIDRTIEQMAQLLAELKSL---LSPOSQNAATGAGNDQTTGVGNAGLNGRKGATG 93  
Db 72 KPNDQS--NIAKLISALIMSLQLMLTNSKKQDTNQEQPDSQAPQNNGLG-----122  
  
QY 94 TTPQSDSQNMLSEMGNNGLDQAITPGCGGGQIGDNPLLKAMKLIARWMDGSDQFGQP 153  
Db 123 -TPSADS-----GGGG-----TPDATGGG-GDTP-----SATGGG 151  
  
QY 154 GTGNNSASGTSSTSSGGSPFNDLSGGKAPSGNSPGNSYSPVSTPSPSTPTSPSPDPS 213  
Db 152 GGTPTATGGGGGGGTPTATGGG---SGTPTATGGGGGVTPOITPOL-----A 200  
  
QY 214 SPTKAAGGSTPVTDHPDPVGSAGIGAGNSVAFPSAGANQTVLHDTITVKAGQVFDKGQT 273  
Db 201 NPRTSG-----TGSVSDTAGS---TEQAGKINNVKDTIKVGAGEVFDGHGAT 245  
  
QY 274 FTAGSELGCGQSENQKPLFILEDGASLKNVTMGDDGADGIHLG-----DAKIDNLHVTN 329  
Db 246 FTADKSMGNDQGENQKPMFELAEATLKNVNLGENEVDDGIHVKAQNAQEVITDINVAQN 305  
  
QY 330 VGEDAITVKNPSAGKSHVEITNSSPEHASDKILQNLADTNLSVDNVKAKDFCTFVRTNG 389  
Db 306 VGEDLITVKEGGAATVNLNKNSSAKGADDKVQLNANTHLKIDNFKADDFGTWRTNG 365  
  
QY 390 GQO-GNWDNLNLSHISAEDCKFSFKVSDSEGLNVNTSDISLGDVYENHY 435  
Db 366 GKQFDDMSIELNGIEANHGKFAVKSDDSLKLATGNIAMTDVKHAY 412

```

; APPLICANT: ITO, Susumu
; APPLICANT: TSUMADORI, Masaki
; TITLE OF INVENTION: Detergent Composition
; FILE REFERENCE: 2173-0116P
; CURRENT APPLICATION NUMBER: US/09/402,668
; CURRENT FILING DATE: 1998-10-08
; PRIOR APPLICATION NUMBER: 9-091142 JAPAN
; PRIOR FILING DATE: 1997-04-09
; PRIOR APPLICATION NUMBER: 9-242736 JAPAN
; PRIOR FILING DATE: 1997-09-08
; PRIOR APPLICATION NUMBER: PCT/US98/01613
; PRIOR FILING DATE: 1998-04-09
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Bacillus sp.
; OTHER INFORMATION: Strain: KSM-P15
US-09-402-668-2

Query Match      15.6%; Score 359.5; DB 3; Length 197;
Best Local Similarity 46.3%; Pred. No. 6.6e-21;
Matches 81; Conservative 29; Mismatches 56; Indels 9; Gaps 4;

QY 253 TVLHDTITVKGQVDFGKQFTTAG-SELGGQSENOKPLFILEDGASLKNVTMGDDGA 311
Db 3 TVVHETIRVPAGQTFDGGQTYVNPNTLGDGSAENOKPLFLEAGASLKNVVGAPAA 62

QY 312 DGIHLYGDAKIDNLHVTNCGEDAITVKPNSAGKKSHVETITNSFEHASDKILOLNADTNL 371
Db 63 DGVHCYGDCTTNVWEDVGEDALTK--SSGT---VNISGAAYKAYDKVQFINAAGTI 117

QY 372 SVDNVKADFGTFVRTNGQQGNWDLNLSHISAEDGKFSFKVSDSE---GLNVNT 423
Db 118 NIRNFRADDIGKLRQNGGTYKVMNVNENCNISKVKDAILRTDSTSTGRIVNT 172

RESULT 6
US-09-198-956-4
; Sequence 4, Application US/09198956
; Patent No. 6165769
; GENERAL INFORMATION:
; APPLICANT: Andersen, Lene N.
; APPLICANT: Schulein, Martin
; APPLICANT: Lange, Niels Erik K.
; APPLICANT: Bjornvad, Mads E.
; APPLICANT: Schnorr, Kirk
; TITLE OF INVENTION: Pectin Degrading Enzymes From Bacillus
; FILE REFERENCE: 5377-200-US
; CURRENT APPLICATION NUMBER: US/09/198,956
; CURRENT FILING DATE: 1998-11-24
; EARLIER APPLICATION NUMBER: 1344/97
; EARLIER FILING DATE: 1997-11-24
; EARLIER APPLICATION NUMBER: 60/067,240
; EARLIER FILING DATE: 1997-12-02
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 221
; TYPE: PRT
; ORGANISM: Bacillus licheniformis
US-09-198-956-4

Query Match      14.8%; Score 341; DB 3; Length 221;
Best Local Similarity 41.7%; Pred. No. 2.3e-19;
Matches 73; Conservative 33; Mismatches 59; Indels 10; Gaps 2;

QY 254 VLHDTITVKGQVDFGKQFTTAGSELGGQSENOKPLFILEDGASLKNVTMGDDGADG 313
Db 31 VVHKTIVVEKQTYDGGKRLIAGPELGDGSGQREDDQKPIFKVEDGATLKNVVLGAPADG 90

QY 314 IHLYGDAKIDNLHVTNCGEDAITVKPNSAGKKSHVETITNSFEHASDKILOLNADTNL 373
Db 91 VHTYGNASINNVWEDVGEDALTVK-----SEGSVTINGGSARLAADKIFQINKASTFTV 145

QY 374 DNVKADFGTFVRTNGQQGNWDLNLSHISAEDGKFSFKVSDSEGLNVNTSDISL 428
Db 146 KNFTADQGGKFIROLGSGSTFKAVVNDICTITNKEAIFRTDS-----STSSVTM 195

RESULT 8
US-09-402-668-10
; Sequence 10, Application US/09402668
; Patent No. 6172030
; GENERAL INFORMATION:
; APPLICANT: WADA, Yasunao
; APPLICANT: KASAI, Miyuki
; APPLICANT: SHIKATA, Shitsuw
; APPLICANT: SUZUMATSU, Atsushi
; APPLICANT: KOIKE, Kenzo
; APPLICANT: HATADA, Yuji
; APPLICANT: KOBAYASHI, Tohru
; APPLICANT: ITO, Susumu
; APPLICANT: TSUMADORI, Masaki
; TITLE OF INVENTION: Detergent Composition
; FILE REFERENCE: 2173-0116P
; CURRENT APPLICATION NUMBER: US/09/402,668
; CURRENT FILING DATE: 1998-10-08

```

```

QY 314 IHLYGDAKIDNLHVTNCGEDAITVKPNSAGKKSHVETITNSFEHASDKILOLNADTNL 373
Db 91 VHTYGNASINNVWEDVGEDALTVK-----SEGSVTINGGSARLAADKIFQINKASTFTV 145

QY 374 DNVKADFGTFVRTNGQQGNWDLNLSHISAEDGKFSFKVSDSEGLNVNTSDISL 428
Db 146 KNFTADQGGKFIROLGSGSTFKAVVNDICTITNKEAIFRTDS-----STSSVTM 195

RESULT 7
US-09-670-141-4
; Sequence 4, Application US/09670141
; Patent No. 6429000
; GENERAL INFORMATION:
; APPLICANT: Andersen, Lene N.
; APPLICANT: Schulein, Martin
; APPLICANT: Lange, Niels Erik K.
; APPLICANT: Bjornvad, Mads E.
; APPLICANT: Schnorr, Kirk
; TITLE OF INVENTION: Pectin Degrading Enzymes From Bacillus
; FILE REFERENCE: 5377-200-US
; CURRENT APPLICATION NUMBER: US/09/670,141
; CURRENT FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: 09/198,956
; PRIOR FILING DATE: 1998-11-24
; PRIOR APPLICATION NUMBER: 1344/97
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/067,240
; PRIOR FILING DATE: 1997-12-02
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 221
; TYPE: PRT
; ORGANISM: Bacillus licheniformis
US-09-670-141-4

Query Match      14.8%; Score 341; DB 4; Length 221;
Best Local Similarity 41.7%; Pred. No. 2.3e-19;
Matches 73; Conservative 33; Mismatches 59; Indels 10; Gaps 2;

QY 254 VLHDTITVKGQVDFGKQFTTAGSELGGQSENOKPLFILEDGASLKNVTMGDDGADG 313
Db 31 VVHKTIVVEKQTYDGGKRLIAGPELGDGSGQREDDQKPIFKVEDGATLKNVVLGAPADG 90

QY 314 IHLYGDAKIDNLHVTNCGEDAITVKPNSAGKKSHVETITNSFEHASDKILOLNADTNL 373
Db 91 VHTYGNASINNVWEDVGEDALTVK-----SEGSVTINGGSARLAADKIFQINKASTFTV 145

QY 374 DNVKADFGTFVRTNGQQGNWDLNLSHISAEDGKFSFKVSDSEGLNVNTSDISL 428
Db 146 KNFTADQGGKFIROLGSGSTFKAVVNDICTITNKEAIFRTDS-----STSSVTM 195

RESULT 8
US-09-402-668-10
; Sequence 10, Application US/09402668
; Patent No. 6172030
; GENERAL INFORMATION:
; APPLICANT: WADA, Yasunao
; APPLICANT: KASAI, Miyuki
; APPLICANT: SHIKATA, Shitsuw
; APPLICANT: SUZUMATSU, Atsushi
; APPLICANT: KOIKE, Kenzo
; APPLICANT: HATADA, Yuji
; APPLICANT: KOBAYASHI, Tohru
; APPLICANT: ITO, Susumu
; APPLICANT: TSUMADORI, Masaki
; TITLE OF INVENTION: Detergent Composition
; FILE REFERENCE: 2173-0116P
; CURRENT APPLICATION NUMBER: US/09/402,668
; CURRENT FILING DATE: 1998-10-08

```

```

; PRIOR APPLICATION NUMBER: 9-091142 JAPAN
; PRIOR FILING DATE: 1997-04-09
; PRIOR APPLICATION NUMBER: 9-242736 JAPAN
; PRIOR FILING DATE: 1997-09-08
; PRIOR APPLICATION NUMBER: PCT/US98/01613
; PRIOR FILING DATE: 1998-04-09
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 10
; LENGTH: 62
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:peptide from
; OTHER INFORMATION: primer
; US-09-402-668-10

Query Match      8.2%; Score 190.5; DB 3; Length 62;
Best Local Similarity 65.6%; Pred. No. 3.3e-08;
Matches 39; Conservative 7; Mismatches 13; Indels 1; Gaps 1;

QY 253 TVLHDTITVKAGVFDGKGTFTAG-SELGSGQSQENKQPLFILEDGASIKNVMTMGDDGA 311
Db 3 TVVHETIRVPAGOTFDGKGOTYVNPNTLGDGSAENKQPIFLEAGASIKNVVIGAPAA 62

RESULT 9
US-09-268-347-36
; Sequence 36, Application US/09268347
; Patent No. 6335182
; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena M.
; TITLE OF INVENTION: RECOMBINANT HAEMOPHILUS INFLUENZAE ADHESIN PROTEINS
; FILE REFERENCE: 1038-860
; CURRENT APPLICATION NUMBER: US/09/268,347
; CURRENT FILING DATE: 1999-03-16
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 36
; LENGTH: 2411
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
; US-09-268-347-36

Query Match      7.9%; Score 181.5; DB 3; Length 2411;
Best Local Similarity 20.6%; Pred. No. 2.6e-05;
Matches 126; Conservative 64; Mismatches 205; Indels 217; Gaps 26;

QY 4 LTUN-----NTSSPGLFQSGGNGLGGHN-----ANSALGOQPIDRQT 43
Db 1182 ITLANGAAAGTDSNGNTISVTVDGISAGNKEITNVKSALKTKYKDTQNTAGATQPA-ANT 1240
QY 44 IEQMAQLLAELKLLSPOSNGNAATGAGGNDQTTGVGNAGLNG-----RKGTAGTTPQ 97
Db 1241 AEVAKQDLVDLTK-----PATGAAGNADAKAPDTTAAVTGDLGKGLGVLSAKKTADETQD 1296
QY 98 SD-----SONLSEMNGNLDQAITPDGQ-----GGQIGNPLLKAMKLKIARM 142
Db 1297 KEFHAAVKANAEVEFVGKNGATVSAKTDNNGKHTVIDVAEKVGDG-----LEKD 1347
QY 143 MCGSQDQFGQPGTGN-----SASGTSSSGSGSPNDLSGKAPSGNSPGNSPVSTFSP 198
Db 1348 TDGKIKLVNDTNGNLLTVDATKGASVA-----KGEFNAVTT---1385
QY 199 PSTPTSTPLDFPSPTKAAGSTPV-TDHPD--PVGSAGICAGNSVAP-----245
Db 1386 DATTAQGTNANERKVVVVGKSGNATATETDKKVVATVGDVAKAINDAATPVKVENDDSAT 1445
QY 246 -----TSAGANQTV-LHDTITTVKAGQVF-----DGKGTFTAGSE-----LGD 282
Db 1446 IDSPPTDGDANDALKAGDTITLAKGNLKVKGKGNITFALANDLSVKSATVSDKLSLGT 1505
QY 283 GGQSQENKQPLFILEDGASL-----KNVTMGDDGADGIHLYG-----DAKIDNLHVTNVGSD 333
```

```

Db 1506 NGKVN-----ITSDTKGLNFAKDSKTDGDA--NIHLNGIASLTTLNSGATTNLGNN 1558
QY 334 AIT-----VKPNSAGKSHVEITNSFEHASDKILOLNADTNL 371
Db 1559 GITDNEKKRAASVKDVLNAGMNVGVKGPASANN-----QVENIDFVATYDVTDFVSGDKOT 1614
QY 372 SVDNVKAKDFG-----TFVRTNGGQOQGN-- 394
Db 1615 TSVTVESKONGKRTVEYKIGAKTSVIKOHNGKLPFGKELKDANNNGVTVTETDGRDSEGNL 1674
QY 395 -----WDLNLSHISAEDGKFSFVK-----DSEGLNVNTSDISLGDVEN 433
Db 1675 VTAKAVIDAVKAGWRVKTGTGANGQNDPATVASGNTVTFADGNGTTAEVTKANDGSITV 1734
QY 434 HYKVPMSANLKV 445
Db 1735 KYNVKVADGLKL 1746

RESULT 10
US-09-077-098A-6
; Sequence 6, Application US/09077098A
; Patent No. 6544519
; GENERAL INFORMATION:
; APPLICANT: TOKUNAGA, Eiji
; SAKAGUCHI, Masaaki
; MATSUO, Kazuo
; HAMADA, Fukuaburo
; TOKIYOSHI, Sachio
; TITLE OF INVENTION: NOVEL POLYPEPTIDE FROM HAEMOPHILUS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 624 Ninth Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/077,098A
; FILING DATE: 19-May-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP97/03222
; FILING DATE: 12-SEP-1997
; APPLICATION NUMBER: JP 27,148/1996
; FILING DATE: 19-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: KORNBAU, Anne M.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: TOKUNAGA=1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2042 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
; US-09-077-098A-6

Query Match      7.5%; Score 174; DB 4; Length 2042;
Best Local Similarity 23.2%; Pred. No. 8e-05;
Matches 114; Conservative 64; Mismatches 189; Indels 124; Gaps 22;
```





Best Local Similarity 32.0%; Pred. No. 3.3e-05;  
Matches 57; Conservative 20; Mismatches 80; Indels 21; Gaps 7;

Qy 18 QSGD---NGLGHNANSALGQPIDRTIEQMAQLLAEI-LKSLSPQSGNAATGAGN 73  
Db 134 QPGNDKNGVGGANGAKGAGGGGGLAEALQETIEQLAQLGGGAGAGGGVGGAGGA 193

Qy 74 DQTTGVNAGGLNGRKTAGT--TPQSDSONMLSENGNGLDQAITPDG--QSGGQIGDN 129  
Db 194 DGGSGAGGAGGANGADGGNGVNGNQANGPQAGDVNGANGAD-----DGSDDQGGTLTVL 248

Qy 130 PLLKAMKLIARWMD-----GQSDQFGQPGCTGNNSASSGTSSSGSP--FNDLSGCK 179  
Db 249 QKLMKILNALVQMMQGGGLGGNQAGGSKGAGNAPASGANPQANQPGSADDSGQ 306

RESULT 13  
US-09-030-270A-7  
; Sequence 7, Application US/09030270A  
; Patent No. 5977060  
; GENERAL INFORMATION:  
; APPLICANT: Zitter, Thomas A.  
; APPLICANT: Wei, Zhong-Min  
; TITLE OF INVENTION: INSECT CONTROL WITH A  
; TITLE OF INVENTION: HYPERSENSITIVE RESPONSE ELICITOR  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP  
; STREET: P.O. Box 1051, Clinton Square  
; CITY: Rochester  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 14603  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/030,270A  
; FILING DATE:  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/039,226  
; FILING DATE: 28-FEB-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Goldman, Michael L.  
; REGISTRATION NUMBER: 30,727  
; REFERENCE/DOCKET NUMBER: 19603/1521  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (716) 263-1304  
; TELEFAX: (716) 263-1600  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 344 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-09-030-270A-7

Query Match 7.2%; Score 165.5; DB 2; Length 344;  
Best Local Similarity 32.0%; Pred. No. 3.3e-05;  
Matches 57; Conservative 20; Mismatches 80; Indels 21; Gaps 7;

Qy 18 QSGD---NGLGHNANSALGQPIDRTIEQMAQLLAEI-LKSLSPQSGNAATGAGN 73  
Db 134 QPGNDKNGVGGANGAKGAGGGGGLAEALQETIEQLAQLGGGAGAGGGVGGAGGA 193

Qy 74 DQTTGVNAGGLNGRKTAGT--TPQSDSONMLSENGNGLDQAITPDG--QSGGQIGDN 129  
Db 194 DGGSGAGGAGGANGADGGNGVNGNQANGPQAGDVNGANGAD-----DGSDDQGGTLTVL 248

Qy 130 PLLKAMKLIARWMD-----GQSDQFGQPGCTGNNSASSGTSSSGSP--FNDLSGCK 179  
Db 249 QKLMKILNALVQMMQGGGLGGNQAGGSKGAGNAPASGANPQANQPGSADDSGQ 306

RESULT 14  
US-08-984-207-7  
; Sequence 7, Application US/08984207  
; Patent No. 6235974  
; GENERAL INFORMATION:  
; APPLICANT: Qiu, Dewen  
; APPLICANT: Wei, Zhong-Min  
; APPLICANT: Bear, Steven V.  
; TITLE OF INVENTION: HYPERSENSITIVE RESPONSE INDUCED  
; TITLE OF INVENTION: RESISTANCE IN PLANTS BY SEED TREATMENT  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP  
; STREET: P.O. Box 1051, Clinton Square  
; CITY: Rochester  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 14603  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/984,207  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/033,230  
; FILING DATE: 05-DEC-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Goldman, Michael L.  
; REGISTRATION NUMBER: 30,727  
; REFERENCE/DOCKET NUMBER: 19603/1201  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (716) 263-1304  
; TELEFAX: (716) 263-1600  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 344 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-984-207-7

Query Match 7.2%; Score 165.5; DB 3; Length 344;  
Best Local Similarity 32.0%; Pred. No. 3.3e-05;  
Matches 57; Conservative 20; Mismatches 80; Indels 21; Gaps 7;

Qy 18 QSGD---NGLGHNANSALGQPIDRTIEQMAQLLAEI-LKSLSPQSGNAATGAGN 73  
Db 134 QPGNDKNGVGGANGAKGAGGGGGLAEALQETIEQLAQLGGGAGAGGGVGGAGGA 193

Qy 74 DQTTGVNAGGLNGRKTAGT--TPQSDSONMLSENGNGLDQAITPDG--QSGGQIGDN 129  
Db 194 DGGSGAGGAGGANGADGGNGVNGNQANGPQAGDVNGANGAD-----DGSDDQGGTLTVL 248

Qy 130 PLLKAMKLIARWMD-----GQSDQFGQPGCTGNNSASSGTSSSGSP--FNDLSGCK 179  
Db 249 QKLMKILNALVQMMQGGGLGGNQAGGSKGAGNAPASGANPQANQPGSADDSGQ 306

RESULT 15  
US-09-013-587-7  
; Sequence 7, Application US/09013587  
; Patent No. 6277814  
; GENERAL INFORMATION:

```

; APPLICANT: Qiu, Dewen
; APPLICANT: Wei, Zhong-Min
; APPLICANT: Beer, Steven V.
; TITLE OF INVENTION: ENHANCEMENT OF GROWTH IN PLANTS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP
; STREET: Clinton Square, P.O. Box 1051
; CITY: Rochester
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 14603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/013,587
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/036,048
; FILING DATE: 27-JAN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Goldman, Michael L.
; REGISTRATION NUMBER: 30,727
; REFERENCE/DOCKET NUMBER: 19603/1501
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (716) 263-1304
; TELEFAX: (716) 263-1600
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 344 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-013-587-7

Query Match
Best Local Similarity 7.2%; Score 165.5; DB 3; Length 344;
Matches 57; Conservative 20; Mismatches 80; Indels 21; Gaps 7;

QY 18 QSGGD---NGLGCHNANSALGQPIDRTIETQMAQLLAEL-LKSLSPQSGNAATGAGN 73
Db 134 QPGGNDKNGVGGANGAKGAGGGQGLAEQLQIQLAQLGGGGAGAGGAGGGVGGAGA 193
QY 74 DQTTGVNAGLNGRKGTAGT--TPQSDSQNMLSEMGNGLDQAIPTDG--QGGGQIGDN 129
Db 194 DGGSGAGGAGGANGADGGNGVNGQANGPQVAGDVNGANGAD-----DQSEDDQGLTGV 248
QY 130 PLLKAMKLILARMWD-----GQSDQFGQPGTGNNSSASGTSSSGGSP--FNDLSGGK 179
Db 249 QKLMKILNALVQMVGQGLGGNGAQGGKAGNAGSPASGANPGANQPGSADDQSSGQ 306

RESULT 16
US-09-086-118-27
; Sequence 27, Application US/09086118
; Patent No. 6583107
; GENERAL INFORMATION:
; APPLICANT: Baby, Ronald J.
; APPLICANT: Beer, Steven V.
; APPLICANT: Wei, Zhong-Min
; TITLE OF INVENTION: FRAGMENTS ELICITING A HYPERSENSITIVE RESPONSE AND USES
; TITLE OF INVENTION: FRAGMENTS ELICITING A HYPERSENSITIVE RESPONSE AND USES
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP
; STREET: Clinton Square, P.O. Box 1051
; CITY: Rochester

```

```

; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 14603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/086,118
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/048,109
; FILING DATE: 30-MAY-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Goldman, Michael L.
; REGISTRATION NUMBER: 30,727
; REFERENCE/DOCKET NUMBER: 19603/1301
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (716) 263-1304
; TELEFAX: (716) 263-1600
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 344 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-086-118-27

Query Match
Best Local Similarity 7.2%; Score 165.5; DB 4; Length 344;
Matches 57; Conservative 20; Mismatches 80; Indels 21; Gaps 7;

QY 18 QSGGD---NGLGCHNANSALGQPIDRTIETQMAQLLAEL-LKSLSPQSGNAATGAGN 73
Db 134 QPGGNDKNGVGGANGAKGAGGGQGLAEQLQIQLAQLGGGGAGAGGAGGGVGGAGA 193
QY 74 DQTTGVNAGLNGRKGTAGT--TPQSDSQNMLSEMGNGLDQAIPTDG--QGGGQIGDN 129
Db 194 DGGSGAGGAGGANGADGGNGVNGQANGPQVAGDVNGANGAD-----DQSEDDQGLTGV 248
QY 130 PLLKAMKLILARMWD-----GQSDQFGQPGTGNNSSASGTSSSGGSP--FNDLSGGK 179
Db 249 QKLMKILNALVQMVGQGLGGNGAQGGKAGNAGSPASGANPGANQPGSADDQSSGQ 306

RESULT 17
US-09-431-614-15
; Sequence 15, Application US/09431614
; Patent No. 6624139
; GENERAL INFORMATION:
; APPLICANT: Wei, Zhong-Min
; APPLICANT: Schading, Richard L.
; TITLE OF INVENTION: HYPERSENSITIVE RESPONSE ELICITOR-INDUCED STRESS
; TITLE OF INVENTION: RESISTANCE
; FILE REFERENCE: 21829/41 (EBC-003)
; CURRENT APPLICATION NUMBER: US/09/431,614
; CURRENT FILING DATE: 1999-11-02
; EARLIER FILING DATE: 1998-11-05
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 15
; LENGTH: 344
; TYPE: PRT
; ORGANISM: Pseudomonas solanacearum
; US-09-431-614-15

Query Match
Best Local Similarity 7.2%; Score 165.5; DB 4; Length 344;
Matches 57; Conservative 20; Mismatches 80; Indels 21; Gaps 7;

```

QY 18 QSGGD---NGLGHNANSALGQPPIDRTQIEQMAQLLAEI-LKSLSPSGNNAATGAGN 73  
Db 134 QPGNDKNGVGGANGAKAGGGGGLAEALQEIQLAQLGGGAGAGAGGGVGGAGGA 193  
QY 74 DQTTGVNAGLNGRGTAGT--TPQSDSONMLSEMGNGLDQAITPDG--QGGGQIGDN 129  
Db 194 DGGGAGGAGGANGADGGVNGQNGPQAGDVNGANGAD-----DGEDQGGTGV 248  
QY 130 PLLKAMKLFIARMW-----GQSDQFGPQGTGNNASSTSSSGSP--FNDSLGGK 179  
Db 249 QKLMKILNALVQMVGGLGGGNAQGGSKGAGNAPASGANPQPGSADQSSSQ 306

## RESULT 18

PCT-US96-08819-7  
; Sequence 7, Application PC/TUS9608819  
; GENERAL INFORMATION:  
; APPLICANT: Cornell Research Foundation, Inc.  
; TITLE OF INVENTION: HYPERSENSITIVE RESPONSE INDUCED  
; TITLE OF INVENTION: RESISTANCE IN PLANTS  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP  
; STREET: Clinton Square, P.O. Box 1051  
; CITY: Rochester  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 14603  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US96/08819  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/475,775  
; FILING DATE: 07-JUN-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Goldman, Michael L.  
; REGISTRATION NUMBER: 30,727  
; REFERENCE/DOCKET NUMBER: 19603/10051  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (716) 263-1304  
; TELEFAX: (716) 263-1600  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 344 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
PCT-US96-08819-7

Query Match 7.2%; Score 165.5; DB 5; Length 344;  
Best Local Similarity 32.0%; Pred. No. 3.3e-05;  
Matches 57; Conservative 20; Mismatches 80; Indels 21; Gaps 7;  
QY 18 QSGGD---NGLGHNANSALGQPPIDRTQIEQMAQLLAEI-LKSLSPSGNNAATGAGN 73  
Db 134 QPGNDKNGVGGANGAKAGGGGGLAEALQEIQLAQLGGGAGAGAGGGVGGAGGA 193  
QY 74 DQTTGVNAGLNGRGTAGT--TPQSDSONMLSEMGNGLDQAITPDG--QGGGQIGDN 129  
Db 194 DGGGAGGAGGANGADGGVNGQNGPQAGDVNGANGAD-----DGEDQGGTGV 248  
QY 130 PLLKAMKLFIARMW-----GQSDQFGPQGTGNNASSTSSSGSP--FNDSLGGK 179  
Db 249 QKLMKILNALVQMVGGLGGGNAQGGSKGAGNAPASGANPQPGSADQSSSQ 306

RESULT 19  
US-09-010-928B-4  
; Sequence 4, Application US/09010928B  
; Patent No. 5994099  
; GENERAL INFORMATION:  
; APPLICANT: Lewis, Randolph V  
; APPLICANT: Hayashi, Cheryl Y  
; TITLE OF INVENTION: EXTREMELY ELASTIC SPIDER SILK PROTEIN AND DNA  
; TITLE OF INVENTION: CODING THEREFOR  
; NUMBER OF SEQUENCES: 29  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH  
; STREET: 8110 GATEHOUSE RD. SUITE 500E  
; CITY: FALLS CHURCH  
; STATE: VIRGINIA  
; COUNTRY: UNITED STATES OF AMERICA  
; ZIP: 22042  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/010,928B  
; FILING DATE: 22-JAN-1998  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Murphy Jr., Gerald M  
; REGISTRATION NUMBER: 28977  
; REFERENCE/DOCKET NUMBER: 1447-109P  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 907 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-010-928B-4

Query Match 6.9%; Score 160.5; DB 2; Length 907;  
Best Local Similarity 21.4%; Pred. No. 0.00031;  
Matches 85; Conservative 35; Mismatches 134; Indels 143; Gaps 14;  
QY 12 SSPGLPQSGDNGLGHNNANSALGQPPIDRTQIEQMAQLLAEI-LKSLSPSGNA--ATG 69  
Db 603 SGPGGYGGG--SGAGGTGPGGAGG-----ACGAGGSGG 634  
QY 70 AGNDQTTGCVNAGLNGRGTAGTTPQSD-----SQNMLSEMGNGLDQ 114  
Db 635 AGSGGAGGSGGAGGSGGSGGTITTEDLITIDGADGPTITSELTISGAGSGPGG 694  
QY 115 A----ITPDGQGGGQIGDNPLKAMKLFIARMWQSDQFGPQGTGNNASSTSSSGG-- 169  
Db 695 AGPGGVGPGSGPGGVGPG-----VSGPFG--VGPGSGPGGVGSGSGPGGV 740  
QY 170 -----SPFNDLSGGKAPSGNSPSGN-----YSPVSTFSP-----PST 201  
Db 741 GPGGVGPGSGSGGVGPGGVGPGSGGFGYGPGESEPGYGPSTGYGSGGYPGCGAGPYG 800  
QY 202 PPSPTSPDPPSPPTKAAGGSTPTVDPHPVGSAGIGAGNSVAFTSAGANQTVLHDTIV 261  
Db 801 PGSPGAGY--PGSPGAYYPSRVPDMVNGIMSAMQSGGFNY-----841  
QY 262 KAGQVFDGKGQFTAGSELGDCGQSENQKPLFILEDGASLKNVTMGDGDGADGHLHYDCAK 321  
Db 842 ---QMPGNMLSQYSSGS-----GTCNPNVNVLMAL 870  
QY 322 IDNLH--VTNVGDAITVKPNSAGKSHVITNSSFH 357  
Db 871 LAALHCLSNHGSSSPAPSPTPAAMSAYSNSVGRMFAY 907

## RESULT 20

Query Match		6.8%;	Score 157.5;	DB 4;	Length 2870;			
Best Local Similarity		20.8%;	Pred. No. 0.0026;					
Matches 103;		Conservative 62;	Mismatches 201;	Indels 129;	Gaps 17;			
QY	3	ILTLN----	NNTSSPGLFQSGGDNGLGHNANSALGOQPIDRQITQEWQAALLAELKSL	58				
Db	1079	IFTMNVLP	PTTTTETPTTSSDDAGGKTGTGATG-----	1115				
QY	59	LSPSQGNATGAGGNDQTGTG	NA-----GGLNGRKGT-----	AGTTPQSDQNMLSE	106			
Db	1116	-----GTG	TSGSGGATTLTSGDAVRSTTSGSGSQSGTSGAGSGGT	TASGSGSGSG	1170			
QY	107	MGNGLDQ----	AITPDQGG-----	GQIGDNP	LLKAMLLKLTARMWDGSDQDFQPGT	155		
Db	1171	TGSDGNSGK	TALNGDGTGSGTATPGSHLGD-----	1203				
QY	156	GNNASGCTSSGSGSPFN-	DLUSGKAPGNS-PSGNYS	SPVSTPSPPTSP	TSPSPLDFPS	213		
Db	1204	GGSTSGSGSDNSGSGVSTK	SSSGSDTSGSSDGANGAFSA	TQAQSTRTTKTR	SSLATV	1263		
QY	214	SPTKAAGGSTPVTDPDPVGS-	AGIGAGNSVATFSAGANOTV	LHDTITVKAGQV	DFQKQG	272		
Db	1264	SPISAAEQAIIDAQADVM	QLAGIMDG-SASNNSLNTSSLLNQ	ISSLP	PAADLVEAQS	1322		
QY	273	TFTAGSELGCGGSE	ENOKPFILEDGASLKXNVMGDDGADGI	HLYGDAKIDN	LHNTV--	330		
Db	1323	LLSNTLKIPGVGNMSS	VDVLKTLQDNLIATNLSLAD	EMAKVI-----	TKLANVNM	TSAQS	1377	
QY	331	-----CEDAI	ITKPNASAKGHVEITNSS	FEHAS-----	DKLT	363		
Db	1378	LNSVLSL	DIALKGSTVYTLGVSSTKS	KDGTAVIFGYV	IASGYTLVSPR	CTLSIYGSTI	1437	
QY	364	QLNADTNLSV-----	DNVKAQDGT	FPRTNGCGQGNWDLN	LSHISAE	DGKFSFVKDSEG	418	
Db	1438	YLTGDT	TRASYKQLDGT	VTVA---DTMLAA	AIQIGMFATNGRTV	QVEQDKIDKESL	VSG	1494
QY	419	LNVN	TSDISLGDVEN	433				
Db	1495	-NIMAT	MSGVDVQS	1508				

Search completed: January 25, 2005, 12:08:36  
Job time : 24 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2005, 12:06:11 ; Search time 56.25 Seconds  
(without alignments)  
159.435 Million cell updates/sec

Title: US-09-879-248-6\_COPY\_116\_140  
Perfect score: 128  
Sequence: 1 ITPDQGGGQIGDNPILLKAMKLIA 25

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 100 summaries

Database : A\_Geneseq\_23Sep04:.\*  
1: Geneseq1980s:.\*  
2: Geneseq1990s:.\*  
3: Geneseq2000s:.\*  
4: Geneseq2001s:.\*  
5: Geneseq2002s:.\*  
6: Geneseq2003as:.\*  
7: Geneseq2003bs:.\*  
8: Geneseq2004s:.\*

pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	128	100.0	447	2	Aaw96260 Hypersens
2	128	100.0	447	3	Aay71094 Erwinia a
3	128	100.0	447	3	Aay84855 A hyperse
4	128	100.0	447	5	Aao22548 Hypersens
5	128	100.0	447	5	Aae18296 Erwinia a
6	128	100.0	447	5	Aae16448 E. amylov
7	128	100.0	447	5	Abb09226 Erwinia a
8	57	44.5	206	2	Aaw29770 Malassezi
9	54	42.2	837	3	Aab12381 N-termina
10	54	42.2	837	5	Aae14656 Murine ST
11	54	42.2	837	6	Abui10480 Mouse STA
12	52	40.6	135	2	Aar56974 Partial P
13	52	40.6	135	2	Aaw11979 Mycobacte
14	52	40.6	207	2	Aaw61476 A. fumiga
15	52	40.6	210	8	Adm57310 Modular a
16	52	40.6	347	8	Adm57298 Modular a
17	49	38.3	110	5	Adk34437 Novel hum
18	49	38.3	164	8	Adj25641 Voitage-g
19	49	38.3	436	8	Adn73559 Thale cre
20	49	38.3	613	6	Abra42419 Human pot
21	48.5	37.9	55	5	ABP02086 Human ORF
22	48.5	37.9	193	8	Adn99445 Novel hum
23	48.5	37.9	365	4	ABG60270 Human ova
24	48.5	37.9	365	5	ABG61741 Novel ova
25	48.5	37.9	392	4	ADM19994 Protein e

26	48	37.5	43	7	ADN59761	Adn59761 Peptide-v
27	48	37.5	238	5	ABP73460	Abp73460 Candida a
28	48	37.5	505	4	ABB58292	Abb58292 Drosophil
29	47	36.7	72	7	ADF58708	Adf58708 Human pol
30	47	36.7	451	4	ABG10570	Abg10570 Novel hum
31	47	36.7	737	7	ADP83524	Adp83524 C. reinha
32	47	36.7	737	7	ADP83526	Adp83526 C. reinha
33	47	36.7	872	4	ABB69476	Abb69476 Drosophil
34	47	36.7	1343	4	AUU35444	Auu35444 Haemophil
35	47	36.7	1343	6	ABU30272	Abu30272 Protein e
36	47	36.7	1479	4	ABB70936	Abb70936 Drosophil
37	46.5	36.3	353	8	ADJ43241	Adj43241 Plant tra
38	46.5	36.3	353	8	ADO62808	Ado62808 Transcrip
39	46.5	36.3	504	4	ABB66600	Abb66600 Drosophil
40	46	35.9	402	3	AAG06450	Aag06450 Arabidops
41	46	35.9	402	6	ADA33693	Ada33693 Acinetoba
42	46	35.9	403	3	AAG43157	Aag43157 Arabidops
43	46	35.9	408	4	AAB79473	Aab79473 Corynebac
44	46	35.9	419	7	ADB70132	Adb70132 C. neofor
45	46	35.9	434	3	AAG06449	Aag06449 Arabidops
46	46	35.9	435	3	AAG43156	Aag43156 Arabidops
47	46	35.9	441	4	AAG92180	Aag92180 C glutami
48	46	35.9	749	2	AAR70233	Aar70233 P. falcip
49	46	35.9	749	2	AAW22479	Aaw22479 Plasmodiu
50	46	35.9	749	3	AAV77901	Aav77901 P. falcip
51	46	35.9	828	4	ABB67048	Abb67048 Drosophil
52	46	35.9	1086	5	AAU76760	Aau76760 Plasmodiu
53	46	35.9	1098	7	ABO71440	Ab071440 Pseudomon
54	46	35.9	1150	4	ABB63215	Abb63215 Drosophil
55	45.5	35.5	72	5	ABP02813	Abp02813 Human ORF
56	45.5	35.5	197	4	AAU63166	Aau63166 Propionib
57	45.5	35.5	197	6	ABM59685	Abm59685. Propionib
58	45	35.2	23	7	ADN59774	Adn59774 Peptide-v
59	45	35.2	41	7	ADN59768	Adn59768 Peptide-v
60	45	35.2	46	7	ADN59780	Adn59780 Peptide-v
61	45	35.2	46	7	ADN59786	Adn59786 Peptide-v
62	45	35.2	113	5	ABP41357	Abp41357 Human ova
63	45	35.2	115	4	AAO13141	Aao13141 Human pol
64	45	35.2	133	7	ADF06199	Adf06199 Bacteriol
65	45	35.2	253	3	AAG43442	Aag43442 Arabidops
66	45	35.2	253	3	AAG06682	Aag06682 Arabidops
67	45	35.2	257	3	AAG43441	Aag43441 Arabidops
68	45	35.2	257	3	AAG06681	Aag06681 Arabidops
69	45	35.2	357	2	AAV17969	Aav17969 X. laevis
70	45	35.2	365	2	AAW53102	Aaw53102 Pig H tra
71	45	35.2	365	2	AAW30630	Aaw30630 Swine alp
72	45	35.2	365	2	AAW97356	Aaw97356 Swine alp
73	45	35.2	365	3	AAV79302	Aav79302 Pig alpha
74	45	35.2	365	5	ABA47995	Aba47995 Swine alp
75	45	35.2	403	4	ABB63509	Abb63509 Drosophil
76	45	35.2	515	2	AAW54236	Aaw54236 Human Lam
77	45	35.2	515	4	AAB90770	Aab90770 Human she
78	45	35.2	532	6	ABJ19005	Abj19005 Pathogen
79	45	35.2	532	6	ABM72976	Abm72976 Staphyloc
80	45	35.2	581	6	ABF65400	Abf65400 Bifidobac
81	45	35.2	664	6	ABU04992	Abu04992 Human exp
82	45	35.2	664	7	ADJ63975	Adj63975 Human pro
83	45	35.2	664	7	ADJ70273	Adj70273 Human hea
84	45	35.2	664	8	ADO13765	Ado13765 Human wil
85	45	35.2	665	7	ADE63973	Ad63973 Rat Prote
86	45	35.2	689	4	ABB64493	Abb64493 Drosophil
87	45	35.2	789	6	ABU21326	Abu21326 Protein e
88	45	35.2	874	6	ABU23121	Abu23121 Protein e
89	45	35.2	1254	4	ABB62006	Abb62006 Pseudomon
90	45	35.2	1403	7	ABO76754	Ab076754 Pseudomon
91	45	35.2	1547	4	ABBS8643	Abbs8643 Drosophil
92	45	35.2	1780	4	ABB71483	Abb71483 Drosophil
93	44.5	34.8	202	4	ABG05876	Abg05876 Novel hum
94	44.5	34.8	928	8	ADO40052	Ado40052 Human IC-
95	44.5	34.8	1151	6	ADA34589	Ada34589 Acinetoba
96	44	34.4	97	4	AAO03289	Aao03289 Human pol
97	44	34.4	110	4	AAO63351	Aao63351 Human pro
98	44	34.4	124	3	AAG28294	Aag28294 Arabidops

99 44 34.4 200 3 AAG28293 Aag28293 Arabidops  
100 44 34.4 203 4 AAU64010 Aau64010 Propionib

## ALIGNMENTS

RESULT 1  
AAW96260  
ID AAW96260 standard; protein; 447 AA.  
XX  
AC AAW96260;  
XX  
DT 14-JUN-1999 (first entry)  
XX  
DE Hypersensitive response eliciting protein (HrpW).  
XX  
KW Hypersensitive response elicitor protein; hypersensitive response; hrpW;  
KW pathogen; infection; crop protection; disease resistance;  
KW pest resistance; transgenic plant; colouration; maturation.  
XX  
OS Erwinia amylovora.  
XX  
PN WO9907208-A1.  
XX  
PD 18-FEB-1999.  
XX  
PF 27-JUL-1998; 98WO-US015547.  
XX  
PR 06-AUG-1997; 97US-0055108P.  
XX  
PA (CORR ) CORNELL RES FOUND INC.  
XX  
PI Kim JF, Beer SV;  
XX  
DR WPI; 1999-167126/14.  
DR N-PSDB; AAX09007.  
XX  
PT New Erwinia amylovora hypersensitive response eliciting gene and protein  
PT - useful for providing transgenic plants and seeds with enhanced growth,  
PT and insect and disease resistance.  
XX  
PS Claim 1; Page 50-51; 54pp; English.  
XX

CC The hypersensitive response eliciting protein (hrp) or polypeptide is  
CC produced as part of an active defense by plants against incompatible  
CC pathogen infections. The hypersensitive response is a rapid localised  
CC necrosis. The hrp protein and gene when used in nucleotide constructs are  
CC useful for providing disease resistance to plants, insect control to  
CC plants, and enhancing plant growth (enhancing fruit size and earlier  
CC colouration and maturation), by direct application of the protein to  
CC plants, or by producing transgenic plants or seeds using the hrp gene  
XX  
SQ Sequence 447 AA;

Query Match 100.0%; Score 128; DB 2; Length 447;  
Best Local Similarity 100.0%; Pred. No. 5.9e-11;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ITPDGGGGQIGDNPLLKAMLKIIA 25  
|||  
DB 116 ITPDGGGGQIGDNPLLKAMLKIIA 140

RESULT 2  
AAV71094  
ID AAV71094 standard; protein; 447 AA.  
XX  
AC AAV71094;  
XX  
DT 08-SEP-2000 (first entry)  
XX  
DE Erwinia amylovora hypersensitive response elicitor #2.

XX  
KW Hypersensitive response elicitor; environmental stress resistance; plant.  
XX  
OS Erwinia amylovora.  
XX  
PN WO200028055-A2.  
XX  
PD 18-MAY-2000.  
XX  
PF 04-NOV-1999; 99WO-US026039.  
XX  
PR 05-NOV-1998; 98US-0107243P.  
XX  
PA (EDEN-) EDEN BIOSCIENCE CORP.  
XX  
PI Wei Z, Schading RL;  
XX  
DR WPI; 2000-376566/32.  
DR N-PSDB; AAD00669.  
XX  
PT Application of a hypersensitive response elicitor protein to plants to  
PT impart stress resistance.  
XX  
PS Disclosure; Page 10-12; 84pp; English.  
XX

CC The patent discloses a method to impart stress resistance to plants by  
CC applying a hypersensitive response elicitor in a non-infectious form to a  
CC plant or seed. The present sequence is a hypersensitive response elicitor  
CC protein from Erwinia amylovora. The protein is heat stable, protease  
CC sensitive and suppressed by inhibitors of plant metabolism. It is used to  
CC impart stress resistance to plants  
XX  
SQ Sequence 447 AA;

Query Match 100.0%; Score 128; DB 3; Length 447;  
Best Local Similarity 100.0%; Pred. No. 5.9e-11;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ITPDGGGGQIGDNPLLKAMLKIIA 25  
|||  
DB 116 ITPDGGGGQIGDNPLLKAMLKIIA 140

RESULT 3  
AAV84855  
ID AAV84855 standard; protein; 447 AA.  
XX  
AC AAV84855;  
XX  
DT 08-AUG-2000 (first entry)  
XX  
DE A hypersensitive response elicitor protein.  
XX  
KW Hypersensitive response; insect control; disease resistance;  
KW hypersensitive response elicitor; plant growth; vegetable; crop;  
KW ornamental plant.  
XX  
OS Erwinia amylovora.  
XX  
PN WO200020452-A2.  
XX  
PD 13-APR-2000.  
XX  
PF 05-OCT-1999; 99WO-US023181.  
XX  
PR 05-OCT-1998; 98US-0103050P.  
XX  
PA (EDEN-) EDEN BIOSCIENCE CORP.  
XX  
PI Wei Z, Fan H, Niggemeyer JL;  
XX  
DR WPI; 2000-303745/26.  
DR N-PSDB; AAA14939.

XX Hypersensitive response elicitor polypeptides useful for imparting  
PT enhanced growth, disease resistance and insect resistance to plants,  
PT especially vegetables and ornamental flowers.  
XX  
XX Disclosure; Page 12-13; 100pp; English.  
XX  
XX The present sequence represents a hypersensitive response elicitor  
CC polypeptide. The specification describes hypersensitive response elicitor  
CC polypeptide fragments, which do not elicit a hypersensitive response.  
CC Instead, the proteins impart disease resistance to plants, enhance plant  
CC growth, and/or control insects. The polypeptide fragments may be used to  
CC these properties to plants. The plants which may be treated in this way  
CC include vegetables, crops and ornamental plants such as alfalfa, rice,  
CC wheat, barley, rye, cotton, sunflower, peanut, corn, potato, sweet  
CC potato, bean, pea, chickory, lettuce, cabbage, brussels sprout,  
CC beet, parsnip, turnip, cauliflower, broccoli, radish, spinach, onion,  
CC garlic, eggplant, pepper, celery, carrot, squash, pumpkin, zucchini,  
CC cucumber, apple, pear, melon, citrus, strawberry, grape, raspberry,  
CC pineapple, soybean, tobacco, tomato, sorghum or sugarcane, Arabidopsis  
CC thaliana, Saintpaulia, petunia, pelargonium, poinsettia, chrysanthemum,  
CC carnation or zinnia  
XX  
XX Sequence 447 AA;  
Query Match 100.0%; Score 128; DB 3; Length 447;  
Best Local Similarity 100.0%; Pred. No. 5.9e-11;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ITPDGGGGGQIGDNPLLKAMKLKIIA 25  
DB 116 ITPDGGGGGQIGDNPLLKAMKLKIIA 140  
RESULT 4  
AAO22548  
ID AAO22548 standard; protein; 447 AA.  
XX  
XX AAO22548;  
XX  
XX 28-OCT-2002 (first entry)  
XX  
XX Hypersensitive response elicitor protein #2 from Erwinia amylovora.  
XX  
XX Desiccation inhibitor; longevity enhancer; desiccation inhibitor;  
KW cutting; ornamental plant; hypersensitive response elicitor protein;  
KW transgenic ornamental plant; early flowering; longevity enhancer; plant;  
KW flower bloom; flower.  
XX  
XX Erwinia amylovora.  
OS  
XX WO200237960-A2.  
PN  
XX 16-MAY-2002.  
PD  
XX  
XX 06-NOV-2001; 2001WO-US043715.  
PF  
XX  
XX 13-NOV-2000; 2000US-0248169P.  
PR  
XX  
XX (EDEN-) EDEN BIOSCIENCE CORP.  
PA  
XX  
XX Wei Z, Leon E, Oviedo A;  
PI  
XX WPI: 2002-575194/61.  
DR  
XX N-PSDB; AAL41133.  
DR  
XX Inhibiting desiccation of cuttings from ornamental plants, by treating  
PT ornamental plants with hypersensitive response elicitor protein, or  
PT expressing heterologous hypersensitive response elicitor protein in  
PT plants.  
XX  
XX Disclosure; Page 12-13; 69pp; English.  
XX

CC The invention relates to a method for inhibiting desiccation of cuttings  
CC from ornamental plants. The method involves treating the cuttings with a  
CC hypersensitive response elicitor protein or polypeptide, or providing a  
CC transgenic ornamental plant or plant seed transformed with a DNA molecule  
CC encoding the hypersensitive response elicitor polypeptide, and growing  
CC the ornamental plant or transgenic ornamental plant produced from the  
CC transgenic ornamental plant seed. The hypersensitive response elicitor  
CC protein or polypeptide is useful for inhibiting desiccation of cuttings  
CC from ornamental plants, for harvesting cuttings from ornamental plants,  
CC for promoting early flowering of ornamental plants, and enhancing the  
CC longevity of flower blooms on ornamental plant cuttings. This sequence  
CC represents a hypersensitive response elicitor protein of the invention  
XX  
XX Sequence 447 AA;  
Query Match 100.0%; Score 128; DB 5; Length 447;  
Best Local Similarity 100.0%; Pred. No. 5.9e-11;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ITPDGGGGGQIGDNPLLKAMKLKIIA 25  
DB 116 ITPDGGGGGQIGDNPLLKAMKLKIIA 140  
RESULT 5  
AAE18296  
ID AAE18296 standard; protein; 447 AA.  
XX  
XX AAE18296;  
XX  
XX 07-MAY-2002 (first entry)  
XX  
XX Erwinia amylovora hypersensitive response elicitor (HRE) #2.  
XX  
XX Hypersensitive response elicitor; HRE; transgenic plant; plant growth;  
KW stress tolerance; disease tolerance; modified flower colour;  
KW insect resistance; herbicide resistance; male sterility.  
XX  
XX Erwinia amylovora.  
OS  
XX WO200195724-A2.  
PN  
XX 20-DEC-2001.  
PD  
XX  
XX 13-JUN-2001; 2001WO-US018955.  
PF  
XX  
XX 15-JUN-2000; 2000US-0211585P.  
PR  
XX (EDEN-) EDEN BIOSCIENCE CORP.  
PA  
XX  
XX Wei Z, Derocher J;  
PI  
XX WPI: 2002-130707/17.  
DR  
XX N-PSDB; AAD29125.  
DR  
XX  
XX Improving effectiveness of transgenic plants by topical application of a  
PT hypersensitive response elicitor protein to the transgenic plant or by  
PT incorporating into the plant a transgene encoding the protein.  
XX  
XX Disclosure; Page 13-14; 86pp; English.  
XX  
XX The invention relates to methods of improving the effectiveness of  
CC transgenic plants which involves either topical application of a  
CC hypersensitive response elicitor (HRE) protein to the transgenic plant or  
CC incorporating into the transgenic plant a transgene encoding HRE. HRE  
CC sequence is used for improving the effectiveness of transgenic plants by  
CC maximising the benefit of transgenic traits associated with a deleterious  
CC effect on growth, stress tolerance, disease or insect resistance,  
CC enhanced growth, herbicide resistance, male sterility, modified flower  
CC colour and biochemically modified plant product in the transgenic plants  
CC or overcoming the deleterious effects. The present sequence is Erwinia  
CC amylovora HRE protein  
XX

SQ Sequence 447 AA;

Query Match 100.0%; Score 128; DB 5; Length 447;  
 Best Local Similarity 100.0%; Pred. No. 5.9e-11;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ITPDGGGGQIGDNPPLKAMKLIA 25  
 116 ITPDGGGGQIGDNPPLKAMKLIA 140

DB 116 ITPDGGGGQIGDNPPLKAMKLIA 140

RESULT 6  
 AAE16448  
 ID AAE16448 standard; protein; 447 AA.  
 AC AAE16448;  
 DT 09-APR-2002 (first entry)  
 DE E. amylovora hypersensitive response elicitor protein #2.  
 KW Hypersensitive response elicitor protein; plant growth; fruit coloration;  
 KW disease resistance; stress resistance; phytotoxin; insect infection;  
 KW plant maturation.  
 OS Erwinia amylovora.

FH Key Location/Qualifiers  
 FT Domain 5..64  
 FT /label= Hypersensitive\_response\_eliciting\_domain\_1  
 FT Region 5..45  
 FT /label= Acidic\_unit  
 FT Region 45..64  
 FT /label= Alpha\_helix  
 FT Domain 103..146  
 FT /label= Hypersensitive\_response\_eliciting\_domain\_2  
 FT Region 103..131  
 FT /label= Acidic\_unit  
 FT Region 131..146  
 FT /label= Alpha\_helix

XX WO200198501-A2.  
 XX 27-DEC-2001.  
 XX 12-JUN-2001; 2001WO-US018620.  
 XX 16-JUN-2000; 2000US-0212211P.  
 XX (EDEN-) EDEN BIOSCIENCE CORP.  
 XX Fan H, Wei Z;  
 XX WPI; 2002-122282/16.  
 XX N-PSDB; AAD27016.  
 XX New hypersensitive response elicitor proteins comprising spaced apart  
 PT domains having an acidic portion linked to an alpha-helix, useful for  
 PT imparting disease or stress resistance, controlling insects or enhancing  
 PT plant growth.  
 XX Disclosure; Page 13-14; 99pp; English.

PS The patent discloses hypersensitive response elicitor proteins and  
 CC nucleotides encoding such proteins. Hypersensitive response elicitor  
 CC proteins comprise an isolated pair or more of spaced apart domains, each  
 CC comprising an acidic portion linked to an alpha-helix and capable of  
 CC eliciting a hypersensitive response in plants. Sequences of the invention  
 CC are used to impart disease resistance to plants. Sequences of the invention  
 CC to control insects and/or to impart stress resistance to plants which  
 CC includes resistance to environmental stresses such as climate, air  
 CC pollution, chemical and nutritional stress. The method of imparting  
 CC disease resistance has the potential for treating previously untreatable

CC diseases, treating diseases systemically and avoiding the use of  
 CC infectious agents or environmentally harmful materials. Hyper- sensitive  
 CC response elicitor sequences are used to enhance plant growth which  
 CC encompasses greater yield, increased in quantity of seeds produced,  
 CC percentage of seeds germinated, plant size and biomass, bigger fruits,  
 CC earlier fruit coloration and plant maturation. They are also used for  
 CC insect control which encompasses preventing direct insect damage to plant  
 CC by feeding injury, interfering with insect larval feeding on the plants,  
 CC preventing insects from colonising host plants and releasing phytotoxins.  
 CC Sequences of the invention also prevent subsequent disease damage to  
 CC plants resulting from insect infection. The present sequence is Erwinia  
 CC amylovora hypersensitive response elicitor protein

SQ Sequence 447 AA;

Query Match 100.0%; Score 128; DB 5; Length 447;  
 Best Local Similarity 100.0%; Pred. No. 5.9e-11;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ITPDGGGGQIGDNPPLKAMKLIA 25  
 116 ITPDGGGGQIGDNPPLKAMKLIA 140

DB 116 ITPDGGGGQIGDNPPLKAMKLIA 140

RESULT 7  
 ABB09226  
 ID ABB09226 standard; protein; 447 AA.  
 AC ABB09226;  
 DT 08-JUL-2002 (first entry)  
 DE Erwinia amylovora hypersensitive response elicitor SEQ ID NO:5.  
 KW Hypersensitive response elicitor; fruit; vegetable; plant; desiccation;  
 KW postharvest disease.  
 OS Erwinia amylovora.  
 PN WO200180639-A2.  
 XX 01-NOV-2001.  
 XX 17-APR-2001; 2001WO-US012468.  
 XX 19-APR-2000; 2000US-0198359P.  
 XX (EDEN-) EDEN BIOSCIENCE CORP.  
 XX Wei Z, Qiu D, Remick D;  
 XX WPI; 2002-041357/05.  
 XX N-PSDB; ABL51711.  
 XX Inhibiting post harvest disease (caused by Penicillium, Botrytis,  
 PT Phytophthora, or Erwinia) or desiccation and enhancing the longevity in a  
 PT fruits or vegetables, using hypersensitive response elicitor proteins or  
 PT nucleic acids.  
 XX Disclosure; Page 11-12; 72pp; English.

PS The present invention describes methods for inhibiting post harvest  
 CC disease or desiccation and enhancing the longevity in a fruits or  
 CC vegetables, using hypersensitive response elicitor proteins or  
 CC polypeptides or nucleic acids (I) derived from pathogens (e.g. Erwinia  
 CC amylovora, E. stewartii, E. chrysanthemi, E. carotovora, Xanthomonas,  
 CC pseudomonas syringae, P. solanacearum, Phytophthora, and Clavibacter).  
 CC (I) has bactericidal activity, and can be used in gene therapy. The  
 CC method can be used for inhibiting post harvest disease (caused by  
 CC Penicillium, Botrytis, Phytophthora, or Erwinia) or desiccation and  
 CC enhancing the longevity in a fruits or vegetables. The method enables  
 CC growers, warehouse packers, shippers and suppliers to process, handle and  
 CC store fruit and vegetables with reduced losses caused by post harvest





Db 802 GEGGSLGSQLK 815

RESULT 10  
AAE14656  
ID AAE14656 standard; protein; 837 AA.  
XX  
AC AAE14656;  
XX  
DT 16-JUL-2002 (first entry)  
XX  
DE Murine STAT6 protein.  
XX  
KW Signal transducer and activator of transcription; STAT6;  
KW drug development; drug discovery; crystal; inflammation; allergy; asthma;  
KW leukaemia; anaemia; neutropaenia; thrombocytopaenia; cancer; obesity;  
KW viral disease; growth retardation; murine.  
XX  
OS Mus sp.

Key Location/Qualifiers  
Domain 1..125  
/note= "Conserved N-terminal domain of the STAT family"

US6312887-B1.  
06-NOV-2001.  
24-APR-2000; 2000US-00556273.  
23-JAN-1998; 98US-00012710.  
(UYRQ) UNIV ROCKEFELLER.  
Vinkemeier U, Moarefi I, Darnell JB, Kuriyan J;  
WPI; 2002-033337/04.  
Identifying compounds that bind to signal transducer and activator of transcription proteins, useful for the production of new drugs.  
Example; Col 63-68; 44pp; English.

The invention relates to methods for detecting compounds that bind to signal transducer and activator of transcription (STAT) proteins for the discovery and development of new drug compounds based on the structural properties of the protein crystal. The methods include: identifying a compound that binds to the N-terminal domain of a STAT protein, identifying a compound that enhances or diminishes the binding of the dimeric STAT proteins to each other and/or their nucleic acid binding site; or identifying a compound that enhances or diminishes the ability of STAT protein dimers to induce the expression of a gene operably under the control of a promoter containing at least two adjacent weak binding sites for STAT protein dimers. The methods are used for identifying new drugs. An antagonist of STAT N-terminal dimeric interactions that inhibits the binding of the STAT dimers to adjacent weak binding sites on a promoter of a gene, could be useful as drugs in the treatment of diseases, e.g. inflammation, allergy, asthma and leukaemias. On the other hand, an agonist of N-terminal dimeric interactions between STAT dimers, can be used as drugs in the treatment of diseases e.g. anaemia, neutropaenia, thrombocytopaenia, cancer, obesity, viral diseases and growth retardation. The present sequence is murine STAT6 protein

Sequence 837 AA;

Query Match 42.2%; Score 54; DB 5; Length 837;  
Best Local Similarity 64.3%; Pred. No. 30;  
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 5 GGGGQIGDNPLK 18  
Db 802 GEGGSLGSQLK 815

RESULT 11  
ABU10480  
ID ABU10480 standard; protein; 837 AA.  
XX  
AC ABU10480;  
XX  
DT 06-AUG-2003 (first entry)  
XX  
DE Mouse STAT6 protein.  
XX  
KW Mouse; signal transducer and activator of transcription; drug design;  
KW drug screening; STAT-STAT dimer interaction; STAT6.  
XX  
OS Mus sp.  
XX  
FH Key Location/Qualifiers  
Region 4..9  
/label= alpha\_helix\_1  
Region 11..18  
/label= alpha\_helix\_2  
Region 24..29  
/label= alpha\_helix\_3  
Region 31..36  
/label= alpha\_helix\_4  
Region 39..43  
/label= alpha\_helix\_5  
Region 50..73  
/label= alpha\_helix\_6  
/note= "Residues 57, 61, 65, 68 and 71 contribute to packing of the coiled-coil"  
Region 75..90  
/label= alpha\_helix\_7  
/note= "Residues 77, 81, 84 and 88 contribute to packing of the coiled-coil"  
Region 93..113  
/label= alpha\_helix\_8  
US2003003563-A1.  
02-JAN-2003.  
19-OCT-2001; 2001US-00045792.  
23-JAN-1998; 98US-00012710.  
24-APR-2000; 2000US-00556273.  
(VINK/) VINKEMEIER U.  
(MOAR/) MOAREFI I.  
(DARN/) DARNELL J E.  
(KURI/) KURIYAN J.  
Vinkemeier U, Moarefi I, Darnell JB, Kuriyan J;  
WPI; 2003-447354/42.  
New crystal having an N-terminal domain of a STAT protein performing X-ray crystallographic studies, useful for screening drugs that enhance or inhibit STAT-STAT dimer interactions.  
Disclosure; Page 33-35; 46pp; English.  
The invention relates to a crystal of an N-terminal domain of signal transducer and activator of transcription (STAT) protein, where the crystal effectively diffracts X-rays for the determination of the atomic coordinates of the N-terminal domain of the STAT protein to a resolution of greater than 5.0 Angstrom. The methods and compositions are useful for the design and screening of drugs that enhance or inhibit STAT-STAT dimer interactions. The present sequence represents the amino acid sequence of mouse STAT6 protein  
Sequence 837 AA;

Query Match 42.2%; Score 54; DB 6; Length 937;  
 Best Local Similarity 64.3%; Pred. No. 30;  
 Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 5 GGGGGGIGDNPPLK 18  
 | : | | | : | | | |  
 Db 802 GEGGSLGSQPLK 815

## RESULT 12

AAR56974  
 ID AAR56974 standard; protein; 135 AA.

XX AC AAR56974;  
 XX DT 28-FEB-1995 (first entry)

DE Partial PMGA 1.3 protein of Mycoplasma gallisepticum.  
 XX KW PMGA; adhesin gene complex; hemagglutinin; conserved sequences; primers;  
 KW probes; amplification; polymerase chain reaction; specific; detection;  
 KW PCR; 1.2; 1.3.

XX OS Mycoplasma gallisepticum.  
 XX FH Key Location/Qualifiers  
 FT Peptide 1. .25  
 FT /label= putative signal peptide  
 FT Protein 26.135  
 FT /label= mature PMGA1.2 protein

XX PN AU9350593-A.  
 XX PD 26-MAY-1994.

XX PF 10-NOV-1993; 93AU-00050593.  
 XX PR 10-NOV-1992; 92AU-00005744.

XX (UYME ) UNIV MELBOURNE.

XX PI Browning GF, Markham PF, Whithear KG, Walker ID, Glew MD;  
 XX DR N-PSDB; AAQ88668.

XX WPI; 1994-209061/26.  
 XX DR N-PSDB; AAQ88668.

XX PT Recombinant DNA constructs for Mycoplasma gallisepticum - for diagnosis,  
 PT treatment and prophylaxis of poultry respiratory disorders.  
 XX PS Example 1; Fig 3; 51pp; English.  
 CC AAR56974 shows the partial amino acid sequence of PMGA1.3 derived from  
 CC the partial DNA sequence of a 10 kb insert from a Mycoplasma  
 CC gallisepticum genomic DNA library, detected by probes based on the T3 and  
 CC C7 peptides. The proteins derived show high homology to PMGA. This  
 CC sequence is the best approximation of the sequence given in the  
 CC specification due to poor print quality. Mycoplasma gallisepticum  
 CC infection in poultry, humans and other animals is of economic importance  
 CC to many industries and it is desirable to produce effective vaccines and  
 CC probes for its detection. The sequences and probes and vaccine vectors of  
 CC the invention can be used for the diagnosis and treatment of Mycoplasma  
 CC gallisepticum infection, and for prophylaxis

XX SQ Sequence 135 AA;

Query Match 40.6%; Score 52; DB 2; Length 135;  
 Best Local Similarity 64.3%; Pred. No. 7.8;  
 Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 TPDGQGGGIGDNP 15  
 | : | | | : | | | |  
 Db 57 TNPNGGGGMGDNP 70

## RESULT 13

AAW11979  
 ID AAW11979 standard; protein; 135 AA.

XX AC AAW11979;  
 XX DT 16-OCT-2003 (revised)  
 DT 23-APR-1997 (first entry)

XX DE Mycobacterium gallisepticum PMGA1.3 adhesin partial sequence.  
 XX KW Adhesin; PMGA; mycoplasma; diagnosis; vaccine; vector;  
 KW respiratory disease; poultry; haemagglutinin.  
 XX OS Mycoplasma gallisepticum; strain S6.

XX FH Key Location/Qualifiers  
 FT Peptide 1. .25  
 FT /label= sig\_peptide  
 FT /note= "the signal peptide shows homology to the PMGA1.2  
 FT signal peptide"

XX PN CA2135330-A.  
 XX PD 11-MAY-1995.

XX PF 08-NOV-1994; 94CA-02135330.

XX PR 10-NOV-1993; 93AU-00050593.  
 PR 20-APR-1994; 94US-00230312.

XX (BROW/) BROWNING G F.

XX PI Browning GF, Markham PF, Whithear KG, Walker ID, Glew MD;  
 XX DR WPI; 1995-241027/32.  
 DR N-PSDB; AAT51531.

XX PT New promoter region from a Mycoplasma gallisepticum adhesin gene - useful  
 PT when coupled to foreign antigen gene, for prodn. of multivalent live  
 PT vaccines, also new probes for detecting Mycoplasma and manipulating its  
 PT genome.  
 XX PS Disclosure; Fig 3; 81pp; English.

XX CC Adhesin PMGA1.2 (AAW11978) and adhesin PMGA1.3 fragment (AAW11979) are  
 CC products of gene sequences (see also AAT51531) isolated from Mycoplasma  
 CC gallisepticum. DNA constructs incorporating the promoter and/or signal  
 CC sequences of the PMGA genes can be used in the prodn. of multivalent live  
 CC vaccines. The signal peptide sequence is utilised where attachment of an  
 CC exogenous antigen gene to the mycoplasma cell membrane is required.  
 CC (Updated on 16-OCT-2003 to standardise OS field)

XX SQ Sequence 135 AA;

Query Match 40.6%; Score 52; DB 2; Length 135;  
 Best Local Similarity 64.3%; Pred. No. 7.8;  
 Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 TPDGQGGGIGDNP 15  
 | : | | | : | | | |  
 Db 57 TNPNGGGGMGDNP 70

## RESULT 14

AAW61476  
 ID AAW61476 standard; protein; 207 AA.

XX AC AAW61476;  
 XX DT 19-OCT-1998 (first entry)

XX DE A. fumigatus allergen rAsp f6 sequence.  
XX KW Aspergillus fumigatus; allergic bronchopulmonary aspergillosis; allergen;  
XX KW rAsp f6; human; ABPA; asthma; cystic fibrosis.  
XX OS Aspergillus fumigatus.  
XX PN WO9828624-A1.  
XX PD 02-JUL-1998.  
XX PF 19-DEC-1997; 97WO-SF000172.  
XX PR 20-DEC-1996; 96SE-00004815.  
XX PA (PHAA ) PHARMACIA & UPJOHN DIAGNOSTICS AB.  
XX PI Cramer R, Hemmann S, Blaser K;  
XX DR WPI: 1998-377789/32.  
XX DR N-PSDB; AAV45115.  
XX PT Method for diagnosing allergic broncho-pulmonary aspergillosis - by  
PT determining if individual carries antibodies reactive with one or more  
PT ABPA-related recombinant allergens.  
XX PS Disclosure; Page 34; 42pp; English.  
XX CC This represents an Aspergillus fumigatus allergen rAsp f6. The allergens  
CC rAsp f6, rAsp f4 and rAsp f8 are non-secreted proteins from A. fumigatus  
CC and are related to allergic bronchopulmonary aspergillosis (ABPA). The  
CC invention provides a method for diagnosing ABPA in a human individual,  
CC characterised by determining if the individual carries antibodies  
CC reactive with one or more ABPA-related recombinant allergens. The method  
CC can be used to detect ABPA in a patient and differentiate from allergic  
CC sensitisation to Aspergillus fumigatus, particularly in patients  
CC suffering from asthma or cystic fibrosis  
XX SQ Sequence 207 AA;  
  
Query Match 40.6%; Score 52; DB 2; Length 207;  
Best Local Similarity 40.9%; Pred. No. 13;  
Matches 9; Conservative 6; Mismatches 7; Indels 0; Gaps 0;  
  
QY 1 ITPDQGGGGQIGDNPLLKAMLK 22  
Db 81 LAPEKSGGKIDQAPVLKAAIE 102  
  
RESULT 15  
ADM57310  
ID ADM57310 standard; protein; 210 AA.  
XX AC ADM57310;  
XX DT 01-JUL-2004 (first entry)  
XX DE Modular antigen transporter molecule protein SEQ ID NO: 16.  
XX KW modular antigen transporter molecule; MAT molecule; immunosuppressive;  
KW antiallergic; antirheumatic; virucide; antibacterial; cytostatic;  
KW translocation module; targeting module; antigen module.  
XX OS Aspergillus fumigatus.  
XX PN EP1408114-A1.  
XX PD 14-APR-2004.  
XX PF 11-OCT-2002; 2002EP-00022774.  
XX PR 11-OCT-2002; 2002EP-00022774.  
XX PR 11-OCT-2002; 2002EP-00022774.

XX PA (BIOV-) BIOVISION AG.  
XX PI Cramer R, Flueckiger S, Lamping N, Daigle I;  
XX DR WPI: 2004-307083/29.  
XX DR N-PSDB; ADM57309.  
XX PT Modular antigen-transporting molecule, useful for treating, preventing  
PT and diagnosing e.g. autoimmune disease, comprises separate translocation,  
XX targeting and antigen modules.  
XX PS Disclosure; Page 48-49; 69pp; German.  
XX CC The present invention relates to a modular antigen-transporting molecule  
CC (MAT) comprising one each of a translocation module, a targeting module  
CC and an antigen module. MAT molecules and their coding sequences are used  
CC for imaging, and as pharmaceutical, vaccinating and diagnostic agents,  
CC for preventing, reducing and/or stimulating the immune response, and for  
CC treatment of autoimmune, allergic, rheumatic, organ rejection, infection  
CC (bacterial, viral or caused by eukaryotic pathogens) and/or malignant  
CC disease. The present sequence is a MAT molecule shown in the  
XX exemplification of the invention.  
XX SQ Sequence 210 AA;  
  
Query Match 40.6%; Score 52; DB 8; Length 210;  
Best Local Similarity 40.9%; Pred. No. 13;  
Matches 9; Conservative 6; Mismatches 7; Indels 0; Gaps 0;  
  
QY 1 ITPDQGGGGQIGDNPLLKAMLK 22  
Db 84 LAPEKSGGKIDQAPVLKAAIE 105  
  
RESULT 16  
ADM57298  
ID ADM57298 standard; protein; 347 AA.  
XX AC ADM57298;  
XX DT 01-JUL-2004 (first entry)  
XX DE Modular antigen transporter molecule protein SEQ ID NO: 4.  
XX KW modular antigen transporter molecule; MAT molecule; immunosuppressive;  
KW antiallergic; antirheumatic; virucide; antibacterial; cytostatic;  
KW translocation module; targeting module; antigen module.  
XX OS Homo sapiens.  
XX OS Human immunodeficiency virus 1.  
XX OS Aspergillus fumigatus.  
XX PN EP1408114-A1.  
XX PD 14-APR-2004.  
XX PF 11-OCT-2002; 2002EP-00022774.  
XX PR 11-OCT-2002; 2002EP-00022774.  
XX PA (BIOV-) BIOVISION AG.  
XX PI Cramer R, Flueckiger S, Lamping N, Daigle I;  
XX DR WPI: 2004-307083/29.  
XX DR N-PSDB; ADM57297.  
XX PT Modular antigen-transporting molecule, useful for treating, preventing  
PT and diagnosing e.g. autoimmune disease, comprises separate translocation,  
XX targeting and antigen modules.  
XX PS Disclosure; Page 35-37; 69pp; German.

XX The present invention relates to a modular antigen-transporting molecule  
 CC (MAT) comprising one each of a translocation module, a targeting module  
 CC and an antigen module. MAT molecules and their coding sequences are used  
 CC for imaging, and as pharmaceutical, vaccinating and diagnostic agents,  
 CC for preventing, reducing and/or stimulating the immune response, and for  
 CC treatment of autoimmune, allergic, rheumatism, organ rejection, infection  
 CC (bacterial, viral or caused by eukaryotic pathogens) and/or malignant  
 CC disease. The present sequence is a MAT molecule shown in the  
 CC exemplification of the invention.  
 XX  
 SQ Sequence 347 AA;  
 Query Match 40.6%; Score 52; DB 8; Length 347;  
 Best Local Similarity 40.9%; Pred. No. 23;  
 Matches 9; Conservative 6; Mismatches 7; Indels 0; Gaps 0;  
 QY 1 ITPDGGGGGQIGDNPVLKMLK 22  
 Db 221 LAPEKSGGKIDQAPVLKAAIE 242  
 RESULT 17  
 ADK34437  
 ID ADK34437 standard; protein; 110 AA.  
 XX  
 AC ADK34437;  
 XX  
 DT 06-MAY-2004 (first entry)  
 XX  
 DE Novel human polypeptide SeqID6519.  
 XX  
 KW antiarthritic; antiparkinsonian; neuroprotective; nootropic;  
 KW immunosuppressive; cytostatic; antipsoriatic; antiinflammatory;  
 KW antibacterial; antiviral; antifungal; antiparasitic; gene therapy;  
 KW arthritis; Parkinson's; Alzheimer's; autoimmune disease; cancer;  
 KW psoriasis; inflammatory bowel disease; infection; bacteria; virus;  
 KW fungus; parasite; human.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 1..110  
 FT /label= OTHER  
 FT /note= "OTHER= All Xaa's in this sequence are unknown  
 FT amino acids or the site of a stop codon within the DNA  
 FT sequence"  
 FT  
 FT  
 PN WO200216439-A2.  
 XX  
 XX 28-FEB-2002.  
 XX  
 XX 05-MAR-2001; 2001WO-US004941.  
 XX  
 XX 07-MAR-2000; 2000US-00519705.  
 XX  
 XX 19-MAY-2000; 2000US-00574454.  
 XX  
 XX (HYSE-) HYSEQ INC.  
 XX  
 XX Tang YT, Liu C, Drmanac RT;  
 XX WPI; 2002-280918/32.  
 XX  
 XX Isolated polynucleotide encoding bone marrow derived polypeptides useful  
 XX for treating, e.g., Parkinson's, Alzheimer's, cancer, arthritis, Crohn's  
 XX disease, and inflammatory bowel disease.  
 XX  
 XX Claim 20; SEQ ID NO 6519; 504pp; English.  
 XX  
 XX This invention relates to a novel isolated polynucleotide comprising a  
 CC nucleotide sequence selected from one of 1680 sequences, a mature protein  
 CC coding portion of them, an active domain of them and their complementary  
 CC sequences. The invention may be useful for the production of compounds

CC with an antiarthritic, antiparkinsonian, neuroprotective, nootropic,  
 CC immunosuppressive, cytostatic, antipsoriatic, antiinflammatory,  
 CC antibacterial, antiviral, antifungal or antiparasitic activity. In  
 CC addition, the disclosed sequences may be useful for gene therapy. The  
 CC polypeptides or their antibodies are useful for treating many diseases  
 CC such as arthritis, Parkinson's, Alzheimer's, autoimmune diseases, cancer,  
 CC psoriasis, inflammatory bowel disease and infections caused by bacteria,  
 CC viruses, fungi or parasites. The present sequence is that of a human  
 CC polypeptide of the invention.  
 XX  
 SQ Sequence 110 AA;  
 Query Match 38.3%; Score 49; DB 5; Length 110;  
 Best Local Similarity 61.1%; Pred. No. 18;  
 Matches 11; Conservative 1; Mismatches 2; Indels 4; Gaps 1;  
 QY 3 PDGGGGGQIG----DNPL 16  
 Db 10 PNGGGGGQIGAHGAPNPL 27  
 RESULT 18  
 ADJ25641  
 ID ADJ25641 standard; protein; 164 AA.  
 XX  
 AC ADJ25641;  
 XX  
 DT 20-MAY-2004 (first entry)  
 XX  
 DE Voltage-gated potassium channel amino acid sequence hkv1.5.  
 XX  
 KW potassium channel; voltage-gated potassium channel; mutant; yeast;  
 KW detection; neuronal integration; cardiac pacemaking; muscle contraction;  
 KW hormone secretion; cell volume regulation; lymphocyte differentiation;  
 KW cell proliferation.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO2004008103-A2.  
 XX  
 XX 22-JAN-2004.  
 XX  
 XX 14-JUL-2003; 2003WO-US022165.  
 XX  
 XX 12-JUL-2002; 2002US-0395272P.  
 XX  
 XX (CARD-) CARDIOME INC.  
 XX  
 XX Fedida D, Steele D;  
 XX WPI; 2004-156549/15.  
 XX  
 XX New potassium channel comprising mutations producing an open voltage-  
 XX gated channel, useful in information transfer and synaptic functions in  
 XX neurons and participates in neuronal integration, cardiac pacemaking or  
 XX muscle contraction.  
 XX  
 XX Disclosure; Fig 4; 36pp; English.  
 XX  
 XX The present invention describes a potassium channel comprising a voltage-  
 XX gated potassium channel which when expressed in a mutant yeast deficient  
 XX in potassium uptake allows the mutant yeast to grow in the presence of  
 XX media with very low potassium concentration, where the voltage-gated  
 XX potassium channel comprises one or more mutations which produces a  
 XX constitutively open voltage-gated potassium channel. Also described: (1)  
 XX a yeast cell comprising a deficiency in potassium uptake and a  
 XX constitutively open voltage-gated potassium channel which allows the  
 XX yeast cell to grow in the presence of media with very low potassium; (2)  
 XX a recombinant nucleic acid molecule comprising a nucleic acid sequence  
 XX encoding the potassium channel and a promoter sequence operably linked to  
 XX the nucleic acid molecule; and (3) a method of identifying compounds that  
 XX modulates activity of voltage-gated ion channels. The potassium channel  
 XX can be used as a detection system for detecting inhibitors and/or

CC activators of normally voltage-gated ion channels. The channels play a  
CC role in information transfer and synaptic functions in neurons and  
CC participate in neuronal integration, cardiac pacemaking, muscle  
CC contraction, hormone secretion, cell volume regulation, lymphocyte  
CC differentiation and cell proliferation. The present invention describes a  
CC voltage-gated potassium channel, which is given in the exemplification of  
CC the present invention.

XX Sequence 164 AA;

Query Match 38.3%; Score 49; DB 8; Length 164;  
Best Local Similarity 45.5%; Pred. No. 28;  
Matches 10; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 3 PDGGGGGQIGDNPLKAMKLI 24  
||| | | | | : | : | : |  
Db 15 PGGGGGGGQGGQAMSLAIRVI 36

RESULT 19  
ADN73559  
ID ADN73559 standard; protein; 436 AA.

XX AC ADN73559;

XX DT 15-JUL-2004 (first entry)

XX DE Thale cress protein repressed in E2Fa/Dpa expressing plants SeqID 1454.

XX KW plant; transgenic; E2Fa/Dpa transcription factor; growth regulator;  
XX KW animal feed product; thale cress; cell wall biosynthesis;  
XX KW nitrogen metabolism; carbon metabolism.

XX OS Arabidopsis thaliana.

XX PN WO2004035798-A2.

XX PD 29-APR-2004.

XX PF 20-OCT-2003; 2003WO-EP011658.

XX PR 18-OCT-2002; 2002EP-00079408.

XX PA (CROP-) CROPDISEGN NV.

XX PI Inze D, De Veylder L, Vlieghe K;

XX DR WPI; 2004-348466/32.

XX DR N-PSDB; ADN73558.

XX PT Altering plant characteristics, useful for producing plants for enzyme or  
XX PT pharmaceutical production comprising modifying in a plant, expression of  
XX PT one or more nucleic acids and/or modifying level or activity of one or  
XX PT more proteins.

XX PS Claim 1; SEQ ID NO 1454; 134pp; English.

XX CC This invention relates to a novel method for altering one or more plant  
XX CC characteristics. Specifically, it refers to identifying genes that are up  
XX CC or down-regulated in transgenic plants overexpressing the heterodimeric  
XX CC E2Fa/Dpa transcription factor of Arabidopsis and using these sequences to  
XX CC alter plant characteristics accordingly. The present invention describes  
XX CC generating transgenic plants for the production of growth regulators,  
XX CC enzymes, therapeutics, pharmaceuticals and animal feed products, where  
XX CC the altered plant characteristics are selected from increased yield or  
XX CC biomass, enhanced survival capacity, stress tolerance, plant architecture  
XX CC or physiology, altered endoreplication, biochemistry, signal  
XX CC transduction, storage lipid mobilisation and/or altered photosynthesis,  
XX CC each relative to the corresponding wild type plants. Accordingly, these  
XX CC sequences can also be useful as positive or negative selectable markers  
XX CC during transformation of cells or tissues. The identified genes play a  
XX CC role in a variety of biological processes such as DNA replication, cell  
XX CC wall biosynthesis, nitrogen and/or carbon metabolism or they function as

CC transcription factors. This polypeptide sequence is thale cress protein  
CC expressed by a gene repressed 1.3 fold or more in plants overexpressing  
CC the E2Fa/Dpa transcription factor, given in an exemplification of the  
CC invention.

XX Sequence 436 AA;

Query Match 38.3%; Score 49; DB 8; Length 436;  
Best Local Similarity 52.6%; Pred. No. 84;  
Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 3 PDGGGGGQIGDNPLKAML 21  
||| | | | | : | : |  
Db 22 PDRPGGGQINSNSVLVALL 40

RESULT 20  
ABR42419  
ID ABR42419 standard; protein; 613 AA.

XX AC ABR42419;

XX DT 11-AUG-2003 (first entry)

XX DE Human potassium channel Kv1.5.

XX KW Human; potassium channel; Kv1.5; neuronal disease; multiple sclerosis;  
XX KW neuroprotective.

XX OS Homo sapiens.

XX PN WO2003035690-A2.

XX PD 01-MAY-2003.

XX PF 25-OCT-2002; 2002WO-GB004821.

XX PR 25-OCT-2001; 2001GB-00025636.

XX PA (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.

XX PI Dolly JO, Akhtar S, Shamotienko O;

XX DR WPI; 2003-421404/39.

XX PT Preparing a K+ channel fusion protein comprising alpha subunits and beta  
XX PT subunits, useful in the manufacture of a medicament for treating multiple  
XX PT sclerosis.

XX PS Disclosure; Fig 4; 93pp; English.

XX CC The present sequence is the protein sequence of human potassium channel  
XX CC Kv1.5. The invention provides oligomeric potassium channel fusion  
XX CC proteins, preferably comprising Kv1.1 and/or Kv1.2 subunits, and methods  
XX CC of producing them in mammalian cells. The fusion proteins can be used to  
XX CC identify compounds that modulate potassium channel activity. Such  
XX CC modulator compounds are useful for the treatment of multiple sclerosis

XX Sequence 613 AA;

Query Match 38.3%; Score 49; DB 6; Length 613;  
Best Local Similarity 45.5%; Pred. No. 1.2e+02;  
Matches 10; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 3 PDGGGGGQIGDNPLKAMKLI 24  
||| | | | | : | : | : |  
Db 381 PGGGGGGGQGGQAMSLAIRVI 402

Search completed: January 25, 2005, 12:12:07  
Job time : 65.25 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2005, 12:08:41 ; Search time 28.3654 Seconds  
(without alignments)  
58.450 Million cell updates/sec

Title: US-09-879-248-6\_COPY\_116\_140  
Perfect score: 128  
Sequence: 1 ITPDGGGGGIGDNPILLKWLKLLIA 25

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 100 summaries

Database : Issued Patents.AA.\*

1: /cgn2\_6/prodata/1/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/prodata/1/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/prodata/1/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/prodata/1/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/prodata/1/iaa/PCTUS\_COMB.pep.\*  
6: /cgn2\_6/prodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	128	100.0	447	3	US-09-120-927-2
2	128	100.0	447	4	US-09-431-614-6
3	57	44.5	206	4	US-09-091-097-6
4	54	42.2	837	3	US-09-012-710-12
5	54	42.2	837	3	US-09-556-273-12
6	51	39.8	409	3	US-09-140-466-3
7	48	37.5	252	4	US-09-248-796A-18831
8	48	37.5	256	4	US-09-270-767-46260
9	47	36.7	379	2	US-08-576-626A-50
10	46	35.9	402	4	US-09-328-352-4980
11	46	35.9	502	4	US-09-270-767-45060
12	46	35.9	749	2	US-08-568-459A-6
13	46	35.9	749	2	US-08-487-826B-6
14	46	35.9	749	3	US-09-210-288-6
15	46	35.9	1098	4	US-09-252-991A-20186
16	45	35.2	82	4	US-09-248-796A-22679
17	45	35.2	133	4	US-09-543-681A-6484
18	45	35.2	230	4	US-09-270-767-39647
19	45	35.2	230	4	US-09-270-767-54864
20	45	35.2	365	3	US-09-151-592-2
21	45	35.2	365	3	US-09-254-077A-10
22	45	35.2	365	4	US-09-443-766-13
23	45	35.2	373	3	US-09-254-077A-12
24	45	35.2	427	4	US-09-270-767-42751
25	45	35.2	515	2	US-08-705-660-46
26	45	35.2	515	3	US-08-989-045-46
27	45	35.2	664	4	US-09-538-092-838
28	45	35.2	1403	4	US-09-252-991A-25500
29	44.5	34.8	1151	4	US-09-328-352-5876
30	44	34.4	335	4	US-09-252-991A-24116
31	44	34.4	399	4	US-09-489-039A-14312
32	44	34.4	690	4	US-09-248-796A-19169
33	44	34.4	816	4	US-09-540-236-3443
34	43.5	34.0	149	3	US-09-227-357-196
35	43.5	34.0	415	3	US-09-100-193-2
36	43	33.6	359	4	US-09-252-991A-18134
37	43	33.6	450	4	US-09-252-991A-27345
38	43	33.6	515	4	US-09-252-991A-32251
39	43	33.6	619	4	US-09-252-991A-32940
40	43	33.6	706	4	US-09-270-767-40594
41	43	33.6	706	4	US-09-270-767-56210
42	43	33.6	975	4	US-09-328-352-4764
43	42.5	33.2	197	4	US-09-270-767-32061
44	42.5	33.2	197	4	US-09-270-767-47278
45	42.5	33.2	262	4	US-09-248-796A-19720
46	42.5	33.2	265	2	US-08-898-779-2
47	42.5	33.2	265	2	US-08-898-779-4
48	42.5	33.2	265	3	US-09-224-502-2
49	42.5	33.2	265	3	US-09-224-502-4
50	42.5	33.2	652	1	US-08-765-081-6
51	42.5	33.2	652	3	US-09-098-082-6
52	42	32.8	67	1	US-08-691-641-7
53	42	32.8	77	3	US-08-825-852-21
54	42	32.8	77	3	US-09-052-888-21
55	42	32.8	77	4	US-09-723-890-21
56	42	32.8	77	4	US-09-723-901-21
57	42	32.8	77	4	US-09-723-547-21
58	42	32.8	77	4	US-09-724-127-21
59	42	32.8	77	4	US-09-723-931-21
60	42	32.8	77	4	US-09-723-873-21
61	42	32.8	77	4	US-09-724-114-21
62	42	32.8	77	4	US-09-723-913-21
63	42	32.8	77	4	US-09-723-912-21
64	42	32.8	77	4	US-09-724-095-21
65	42	32.8	77	4	US-09-724-157-21
66	42	32.8	77	4	US-09-724-062-21
67	42	32.8	77	4	US-09-724-065-21
68	42	32.8	86	4	US-09-270-767-41133
69	42	32.8	86	4	US-09-270-767-56349
70	42	32.8	88	4	US-09-270-767-37010
71	42	32.8	88	4	US-09-270-767-52227
72	42	32.8	122	2	US-08-820-170A-1
73	42	32.8	122	3	US-09-055-699-1
74	42	32.8	122	3	US-09-273-565-1
75	42	32.8	122	3	US-09-565-538-1
76	42	32.8	122	3	US-09-661-468-1
77	42	32.8	122	4	US-09-976-165-1
78	42	32.8	198	4	US-09-710-279-2792
79	42	32.8	219	2	US-08-625-377-2
80	42	32.8	219	3	US-09-204-841-2
81	42	32.8	233	2	US-08-928-692-31
82	42	32.8	233	3	US-09-339-972-31
83	42	32.8	261	4	US-09-252-991A-28545
84	42	32.8	292	3	US-08-555-568B-19
85	42	32.8	292	3	US-09-519-223-19
86	42	32.8	292	4	US-09-927-180-19
87	42	32.8	298	4	US-09-252-991A-32302
88	42	32.8	326	3	US-09-721-383-4
89	42	32.8	326	3	US-09-721-137-4
90	42	32.8	326	4	US-09-721-251-4
91	42	32.8	326	4	US-10-114-764-4
92	42	32.8	333	4	US-09-489-039A-8678
93	42	32.8	383	4	US-09-252-991A-22283
94	42	32.8	434	4	US-09-248-796A-18389
95	42	32.8	453	4	US-09-252-991A-31634
96	42	32.8	479	4	US-09-489-039A-7208
97	42	32.8	618	3	US-09-134-001C-3169
98	42	32.8	636	3	US-08-737-752A-2
99	42	32.8	636	4	US-09-807-063-5
100	42	32.8	657	1	US-08-264-534-34

## ALIGNMENTS

RESULT 1  
US-09-120-927-2  
; Sequence 2, Application US/09120927  
; Patent No. 6262018  
; GENERAL INFORMATION:  
; APPLICANT: Kim, Jihyun Francis  
; APPLICANT: Beer, Steven V.  
; TITLE OF INVENTION: HYPERSENSITIVE RESPONSE ELICITOR FROM  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Nixon, Hartgrave, Devans & Doyle LLP  
; CITY: Rochester  
; STREET: P.O. Box 1051, Clinton Square  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 14603  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/120,927  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/055,108  
; FILING DATE: 06-AUG-1977  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Goldman, Michael L.  
; REGISTRATION NUMBER: 30,727  
; REFERENCE/DOCKET NUMBER: 19603/1581  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (716) 263-1304  
; TELEFAX: (716) 263-1600  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 447 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-120-927-2

Query Match 100.0%; Score 128; DB 3; Length 447;  
Best Local Similarity 100.0%; Pred. No. 2.5e-11;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ITPDGGGGGIGDNPPLKMLKLLIA 25  
Db 116 ITPDGGGGGIGDNPPLKMLKLLIA 140

RESULT 2  
US-09-431-614-6  
; Sequence 6, Application US/09431614  
; Patent No. 6624139  
; GENERAL INFORMATION:  
; APPLICANT: Wei, Zhong-Min  
; APPLICANT: Schading, Richard L.  
; TITLE OF INVENTION: HYPERSENSITIVE RESPONSE ELICITOR-INDUCED STRESS  
; FILE REFERENCE: 21829/41 (EBC-003)  
; CURRENT APPLICATION NUMBER: US/09/431,614  
; CURRENT FILING DATE: 1999-11-02  
; EARLIER APPLICATION NUMBER: 60/107,243  
; EARLIER FILING DATE: 1998-11-05

; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 6  
; LENGTH: 447  
; TYPE: PRT  
; ORGANISM: Erwinia amylovora  
US-09-431-614-6

Query Match 100.0%; Score 128; DB 4; Length 447;  
Best Local Similarity 100.0%; Pred. No. 2.5e-11;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ITPDGGGGGIGDNPPLKMLKLLIA 25  
Db 116 ITPDGGGGGIGDNPPLKMLKLLIA 140

RESULT 3  
US-09-091-097-6  
; Sequence 6, Application US/09091097  
; Patent No. 6432407  
; GENERAL INFORMATION:  
; APPLICANT: TAKESAKO, KAZUTOH  
; APPLICANT: OKADO, TAKASHI  
; APPLICANT: YAGIHARA, TOMOKO  
; APPLICANT: KURODA, MASANOBU  
; APPLICANT: ONISHI, YOSHIMI  
; APPLICANT: KATO, IKUNOSHIN  
; APPLICANT: AKIYAMA, KAZUO  
; APPLICANT: YASUEDA, HIROSHI  
; APPLICANT: YAMAGUCHI, HIDEYO  
; TITLE OF INVENTION: ANTIGENIC PROTEIN ORIGINATING IN  
; TITLE OF INVENTION: MALASSEZIA  
; NUMBER OF SEQUENCES: 58  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH, LLP  
; STREET: PO BOX 747  
; CITY: FALLS CHURCH  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22040-0747  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/091,097  
; FILING DATE:  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: WEINER, MARC S.  
; REGISTRATION NUMBER: 32,181  
; REFERENCE/DOCKET NUMBER: 1422-0346P  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703-205-8000  
; TELEFAX: 703-205-8050  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 206 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-091-097-6

Query Match 44.5%; Score 57; DB 4; Length 206;  
Best Local Similarity 45.0%; Pred. No. 0.6;  
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;  
QY 1 ITPDGGGGGIGDNPPLKMLK 20  
Db 86 LAFQSEGGGQLNDGPLKQAI 105



```

, COUNTRY: USA
, ZIP: 07601
, COMPUTER READABLE FORM:
, MEDIUM TYPE: Floppy disk
, COMPUTER: IBM PC compatible
, OPERATING SYSTEM: PC-DOS/MS-DOS
, SOFTWARE: PatentIn Release #1.0, Version #1.30
, CURRENT APPLICATION DATA:
, APPLICATION NUMBER: US/09/556,273
, FILING DATE:
, CLASSIFICATION:
, PRIOR APPLICATION DATA:
, APPLICATION NUMBER: 09/012,710
, FILING DATE:
, ATTORNEY/AGENT INFORMATION:
, NAME: Jackson Esq., David A.
, REGISTRATION NUMBER: 26,742
, REFERENCE/DOCKET NUMBER: 600-1-194
, TELECOMMUNICATION INFORMATION:
, TELEPHONE: 201-487-5800
, TELEFAX: 201-343-1684
, TELEX: 133521
, INFORMATION FOR SEQ ID NO: 12:
, SEQUENCE CHARACTERISTICS:
, LENGTH: 837 amino acids
, TYPE: amino acid
, STRANDEDNESS: single
, TOPOLOGY: linear
, MOLECULE TYPE: protein
, HYPOTHETICAL: NO
, UNP-09-556-273-12
,
Query Match 42.2%; Score 54; DB 3; Len
Best Local Similarity 64.3%; Pred. No. 8.1;
Matches 9; Conservative 2; Mismatches 3;
,
QY 5 GGGGGGIGDNPPLK 18
,
DB 802 GGGGSLGSGQLLK 815
,
RESULT 6
US-09-140-466-3
, Sequence 3, Application US/09140466
, Patent No. 6286160
, GENERAL INFORMATION:
, APPLICANT: CLOUGH, BARBARA
, APPLICANT: PREISER, PETER
, APPLICANT: WILSON, ROBERT
, TITLE OF INVENTION: AN EF-TU PROTEIN ENCODED ON THE
, TITLE OF INVENTION: MALARIA PARASITE AND PROTEIN SYN
, TITLE OF INVENTION: EFFECTIVE AS ANTI-MALARIAL COMPO
, FILE REFERENCE: N68837B GCM PJC DP
, CURRENT APPLICATION NUMBER: US/09/140,466
, CURRENT FILING DATE: 1998-08-26
, EARLIER APPLICATION NUMBER: US 60/056,246
, EARLIER FILING DATE: 1997-08-28
, NUMBER OF SEQ ID NOS: 14
, SOFTWARE: PatentIn Ver. 2.1
, SEQ ID NO 3
, LENGTH: 409
, TYPE: PRT
, ORGANISM: Anacystis nidulans
, UNP-09-140-466-3
,
Query Match 39.8%; Score 51; DB 3; Len
Best Local Similarity 57.1%; Pred. No. 10;
Matches 12; Conservative 3; Mismatches 4;
,
QY 6 QGG--GGIDNPPLKMKLI 24
,
DB 183 QGGASGGKDNPVDPKILKM 203

```

```
RESULT 7
US-09-248-796A-18831
; Sequence 18831, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 18831
; LENGTH: 252
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-18831

Query Match      37.5%; Score 48; DB 4; Length 252;
Best Local Similarity 43.5%; Pred. No. 17;
Matches 10; Conservative 4; Mismatches 9; Indels 9; Gaps 0;

Qy      2 TPDGQGGGQIGDNPLKAMKLKLI 24
Db      64 TTDGNGGGGIGIVPTLQIVATV 86

RESULT 8
US-09-270-767-46260
; Sequence 46260, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 46260
; LENGTH: 256
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-46260

Query Match      37.5%; Score 48; DB 4; Length 256;
Best Local Similarity 47.6%; Pred. No. 18;
Matches 10; Conservative 2; Mismatches 9; Indels 9; Gaps 0;

Qy      4 DGGGGGQIGDNPLKAMKLKLI 24
Db      40 DGDGGGGGGGVEVLKAKMSLL 60

RESULT 9
US-08-576-626A-50
; Sequence 50, Application US/08576626A
; Patent No. 5998194
; GENERAL INFORMATION:
; APPLICANT: Summers, R.G.
; APPLICANT: Katz, L.
; APPLICANT: Donadio, S.
; APPLICANT: Staver, M.J.
; TITLE OF INVENTION: POLYKETIDE-ASSOCIATED SUGAR
; TITLE OF INVENTION: BIOSYNTHESIS GENES
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Abbott Laboratories

US-08-576-626A-50
; Sequence 46260, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 4980
; LENGTH: 402
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-4980

Query Match      35.9%; Score 46; DB 4; Length 402;
Best Local Similarity 39.1%; Pred. No. 59;
Matches 9; Conservative 5; Mismatches 9; Indels 9; Gaps 0;

Qy      1 ITPDGGGGGQIGDNPLKAMKLKLI 23
Db      147 VAPDYQGLGTIGAHPPYNAKSQL 169

RESULT 11
US-09-270-767-45060
; Sequence 45060, Application US/09270767
```

Patent No. 6703491  
; GENERAL INFORMATION:  
; APPLICANT: Homburger et al.  
; TITLE OF INVENTION: Nucleic acids and proteins of *Drosophila melanogaster*  
; FILE REFERENCE: File Reference: 7326-094  
; CURRENT APPLICATION NUMBER: US/09/270,767  
; CURRENT FILING DATE: 1999-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 45060  
; LENGTH: 502  
; TYPE: PRT  
; ORGANISM: *Drosophila melanogaster*  
; FEATURE:  
; OTHER INFORMATION: Xaa means any amino acid  
US-09-270-767-45060

Query Match 35.9%; Score 46; DB 4; Length 502;  
Best Local Similarity 56.2%; Pred. No. 75;  
Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 3 PDGQGGGQIGDNPPLK 18  
Db 405 PDGCGGAGGDRPSIK 420  
||| ||| ||| :|

RESULT 12  
US-08-568-459A-6  
; Sequence 6, Application US/08568459A  
; Patent No. 5849306  
; GENERAL INFORMATION:  
; APPLICANT: Sim, Kim L.  
; APPLICANT: Chitnis, Chetan  
; APPLICANT: Miller, Louis H.  
; APPLICANT: Peterson, David S.  
; APPLICANT: Su, Xin-zhaun  
; APPLICANT: Wellem, Thomas E.  
; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX  
; TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS  
; NUMBER OF SEQUENCES: 37  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Knobbe Martens Olson & Bear  
; STREET: 620 Newport Center Drive 16th Floor  
; CITY: Newport Beach  
; STATE: California  
; COUNTRY: US  
; ZIP: 92660  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/568,459A  
; FILING DATE: 07-DEC-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Israel, Ned  
; REGISTRATION NUMBER: 29,655  
; REFERENCE/DOCKET NUMBER: NIH21.001CPI  
; TELEPHONE: (619) 235-8550  
; TELEFAX: (619) 235-0176  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 749 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; ORIGINAL SOURCE:  
; ORGANISM: *Plasmodium falciparum*

US-08-568-459A-6

Query Match 35.9%; Score 46; DB 2; Length 749;  
Best Local Similarity 47.1%; Pred. No. 1.2e+02;  
Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 2 TPDGQGGGQIGDNPPLK 18  
Db 432 TPNSCGGNYGDRQISK 448  
||| ||| ||| :|

RESULT 13  
US-08-487-826B-6  
; Sequence 6, Application US/08487826B  
; Patent No. 5993827  
; GENERAL INFORMATION:  
; APPLICANT: Sim, Kim L.  
; APPLICANT: Chitnis, Chetan  
; APPLICANT: Miller, Louis H.  
; APPLICANT: Peterson, David S.  
; APPLICANT: Su, Xin-zhaun  
; APPLICANT: Wellem, Thomas E.  
; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX  
; TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS  
; NUMBER OF SEQUENCES: 45  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Knobbe Martens Olson & Bear  
; STREET: 620 Newport Center Drive 16th Floor  
; CITY: Newport Beach  
; STATE: California  
; COUNTRY: US  
; ZIP: 92660  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/487,826B  
; FILING DATE: 10-SEP-1993  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Israel, Ned  
; REGISTRATION NUMBER: 29,655  
; REFERENCE/DOCKET NUMBER: NIH21.001CPI  
; TELEPHONE: (619) 235-8550  
; TELEFAX: (619) 235-0176  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 749 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; ORIGINAL SOURCE:  
; ORGANISM: *Plasmodium falciparum*

US-08-487-826B-6

Query Match 35.9%; Score 46; DB 2; Length 749;  
Best Local Similarity 47.1%; Pred. No. 1.2e+02;  
Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 2 TPDGQGGGQIGDNPPLK 18  
Db 432 TPNSCGGNYGDRQISK 448  
||| ||| ||| :|

RESULT 14  
US-09-210-288-6  
; Sequence 6, Application US/09210288  
; Patent No. 6392026

```
; GENERAL INFORMATION:
; APPLICANT: Sim, Kim L.
; APPLICANT: Chitnis, Chetan
; APPLICANT: Miller, Louis H.
; APPLICANT: Peterson, David S.
; APPLICANT: Su, Xin-zhaun
; APPLICANT: Wellens, Thomas E.
; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
; TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Knobbe Martens Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: California
; COUNTRY: US
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/09/210,288
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Fuller, Michael
; REGISTRATION NUMBER: 36,516
; REFERENCE/DOCKET NUMBER: NIH121.1FWDV1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 749 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Plasmodium falciparum
; US-09-210-288-6

Query Match 35.9%; Score 46; DB 3; Length 749;
Best Local Similarity 47.1%; Pred. No. 1.2e+02;
Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 2 TPDGGGGGQIGDNPILK 18
||:|||||:|
Db 432 TPNSSGGNGYGRQISK 448

RESULT 15
US-09-252-991A-20186
; Sequence 20186, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 20186
; LENGTH: 1098
; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-20186

Query Match 35.9%; Score 46; DB 4; Length 1098;
Best Local Similarity 47.6%; Pred. No. 1.8e+02;
Matches 10; Conservative 2; Mismatches 5; Indels 4; Gaps 1;

QY 1 ITPDGGGGGQ-----IGDNPIL 17
:|||||:|
Db 594 VQPDGGGAORRRHRRAGEGPHL 614

RESULT 16
US-09-248-796A-22679
; Sequence 22679, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 22679
; LENGTH: 82
; TYPE: PRT
; ORGANISM: Candida albicans
; US-09-248-796A-22679

Query Match 35.2%; Score 45; DB 4; Length 82;
Best Local Similarity 45.0%; Pred. No. 14;
Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 5 GQGGGQIGDNPILKAMKLKI 24
||:|||||:|
Db 27 GNGNGSGGNTILQILQLI 46

RESULT 17
US-09-543-681A-6484
; Sequence 6484, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 6484
; LENGTH: 133
; TYPE: PRT
; ORGANISM: Proteus mirabilis
; US-09-543-681A-6484

Query Match 35.2%; Score 45; DB 4; Length 133;
Best Local Similarity 44.4%; Pred. No. 24;
Matches 8; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 2 TPDGGGGQIGDNPILKA 19
:|||||:|
Db 67 ITEVGEGVGEQPLIEA 84

RESULT 18
```

```

US-09-270-767-39647
; Sequence 39647, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 39647
; LENGTH: 230
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-39647

Query Match      35.2%; Score 45; DB 4; Length 230;
Best Local Similarity 33.3%; Pred.No. 45;
Matches       7; Conservative   8; Mismatches    6; Indels    0; Gaps    0;

QY      4     DGCGGGGIGDNPCLKMLKLI 24
          |||:::||::||: : ::
DB      110   DGSGGELDSPLQQPLQVV 130

RESULT 19
US-09-270-767-54864
; Sequence 54864, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 54864
; LENGTH: 230
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-54864

Query Match      35.2%; Score 45; DB 4; Length 230;
Best Local Similarity 33.3%; Pred.No. 45;
Matches       7; Conservative   8; Mismatches    6; Indels    0; Gaps    0;

QY      4     DGCGGGGIGDNPCLKMLKLI 24
          |||:::||::||: : ::
DB      110   DGSGGELDSPLQQPLQVV 130

RESULT 20
US-09-151-592-2
; Sequence 2, Application US/09151592
; Patent No. 6355859
; GENERAL INFORMATION:
; APPLICANT: Bosworth, Brad
; APPLICANT: Ridpath, Julia
; APPLICANT: Wiseman, Barry
; TITLE OF INVENTION: INTERACTIONS BETWEEN GENOTYPE AND DIET IN SWINE THAT
; PREVENT E. COLI ASSOCIATED INTESTINAL DISEASE
; FILE REFERENCE: 21419-90119
; CURRENT APPLICATION NUMBER: US/09/151,592
; CURRENT FILING DATE: 1999-09-18
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 365
; TYPE: PRT
; ORGANISM: Swine
; FEATURE:
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2005, 12:09:52 ; Search time 46.6346 Seconds  
(without alignments)  
193.681 Million cell updates/sec

Title: US-09-879-248-6\_COPY\_116\_140

Perfect score: 128

Sequence: 1 ITPDGGGGGIGDNPILLKAMKLIA 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1608061 seqs, 361289386 residues

Total number of hits satisfying chosen parameters: 1608061

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

#### Database :

Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB\_PUB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.\*
- 19: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 20: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	128	100.0	447	9	US-09-835-684-5
2	128	100.0	447	9	US-09-880-371-5
3	128	100.0	447	9	US-09-879-248-6
4	128	100.0	447	14	US-10-010-390-5
5	128	100.0	447	15	US-10-441-736-6
6	128	100.0	447	17	US-10-847-143-5
7	57	44.5	206	14	US-10-109-670-10
8	54	42.2	837	14	US-10-045-792-12
9	51	39.8	409	9	US-09-845-335-3
10	50	39.1	78	15	US-10-424-599-214915
11	50	39.1	177	16	US-10-437-963-149121
12	49	38.3	149	17	US-10-425-115-231795
13	49	38.3	483	14	US-10-369-493-3880
					Sequence 3880, Ap
					Sequence 5, Appli
					Sequence 5, Appli
					Sequence 6, Appli
					Sequence 5, Appli
					Sequence 5, Appli
					Sequence 10, Appl
					Sequence 12, Appl
					Sequence 3, Appli
					Sequence 214915,
					Sequence 149121,
					Sequence 231795,
					Sequence 3880, Ap

365	9	US-09-908-711-100	Sequence 100, App
43	14	US-10-269-806-163	Sequence 163, App
111	16	US-10-437-963-124136	Sequence 124136,
238	14	US-10-032-585-7297	Sequence 7297, Ap
207	17	US-10-425-115-301902	Sequence 301902,
81	17	US-10-425-115-347563	Sequence 347563,
102	15	US-10-424-598-216781	Sequence 216781,
134	17	US-10-425-115-351040	Sequence 351040,
213	17	US-10-425-115-210287	Sequence 210287,
435	14	US-10-369-493-6514	Sequence 6514, Ap
1343	9	US-09-815-242-11037	Sequence 11037, A
1343	15	US-10-282-122A-58196	Sequence 58196, A
353	15	US-10-374-780A-1704	Sequence 1704, Ap
353	16	US-10-437-963-194124	Sequence 194124,
354	16	US-10-437-963-195060	Sequence 195060,
354	16	US-10-437-963-195060	Sequence 195060,
408	16	US-10-781-014-462	Sequence 462, App
419	15	US-10-320-797-3176	Sequence 3176, Ap
441	9	US-09-738-626-5934	Sequence 5934, Ap
450	17	US-10-425-115-332466	Sequence 332466,
469	15	US-10-425-114-57241	Sequence 57241, A
749	13	US-10-153-273-6	Sequence 6, Appli
1086	9	US-09-924-154-15	Sequence 15, Appl
113	15	US-10-424-598-183658	Sequence 183658,
161	14	US-10-369-493-5035	Sequence 5035, Ap
880	16	US-10-437-963-119573	Sequence 119573,
23	14	US-10-269-806-176	Sequence 176, App
41	14	US-10-269-806-170	Sequence 170, App
46	14	US-10-269-806-182	Sequence 182, App
46	14	US-10-269-806-188	Sequence 188, App
69	16	US-10-437-963-128768	Sequence 128768,
79	17	US-10-425-115-333561	Sequence 333561,
104	17	US-10-425-115-239179	Sequence 239179,
108	17	US-10-425-115-220516	Sequence 220516,
113	15	US-10-264-049-2489	Sequence 2489, Ap
119	16	US-10-437-963-180990	Sequence 180990,
147	17	US-10-425-115-208583	Sequence 208583,
193	16	US-10-437-963-183514	Sequence 183514,
340	16	US-10-437-963-178172	Sequence 178172,
365	9	US-09-051-034A-4	Sequence 4, Appli
365	9	US-09-254-077A-10	Sequence 10, Appl
365	9	US-09-844-268-13	Sequence 13, Appl
365	9	US-09-844-705-13	Sequence 13, Appl
373	9	US-09-254-077A-12	Sequence 12, Appl
476	16	US-10-437-963-164622	Sequence 164622,
476	16	US-10-437-963-202306	Sequence 202306,
515	17	US-09-315-355-45	Sequence 46, Appl
518	17	US-10-425-115-239190	Sequence 239190,
546	15	US-10-425-114-60000	Sequence 60000, A
593	16	US-10-437-963-108680	Sequence 108680,
664	16	US-10-408-765A-2079	Sequence 2079, Ap
664	17	US-10-473-127-1658	Sequence 1658, Ap
692	16	US-10-437-963-161047	Sequence 161047,
789	15	US-10-282-122A-49250	Sequence 49250, A
874	15	US-10-282-122A-51045	Sequence 51045, A
159	17	US-10-425-115-305395	Sequence 305395,
453	15	US-10-425-114-49175	Sequence 49175, A
585	15	US-10-425-114-71739	Sequence 71739, A
655	17	US-10-425-115-305399	Sequence 305399,
701	14	US-10-156-761-12041	Sequence 12041, A
936	17	US-10-425-115-205479	Sequence 205479,
942	15	US-10-425-114-68187	Sequence 68187, A
45	17	US-10-425-115-324889	Sequence 324889,
51	17	US-10-425-115-235183	Sequence 235183,
95	15	US-10-425-114-52465	Sequence 52465, A
163	16	US-10-425-115-233374	Sequence 233374,
163	16	US-10-437-963-166199	Sequence 166199,
235	17	US-10-425-115-249433	Sequence 249433,
252	16	US-10-437-963-146435	Sequence 146435,
284	17	US-10-425-115-208441	Sequence 208441,
285	14	US-10-369-493-19127	Sequence 19127, A
285	16	US-10-437-963-185712	Sequence 185712,
389	14	US-10-369-493-4771	Sequence 4771, Ap
392	14	US-10-369-493-7531	Sequence 7531, Ap

Sequence 1004, Ap  
 Sequence 4993, Ap  
 Sequence 166433, A  
 Sequence 44941, A  
 Sequence 1123, Ap  
 Sequence 134108, A  
 Sequence 63166, A  
 Sequence 78522, A  
 Sequence 14378, A  
 Sequence 184559, A  
 Sequence 339149, A  
 Sequence 194061, A  
 Sequence 209361, A  
 Sequence 257, App

ALIGNMENTS

RESULT 1  
 US-09-835-684-5  
 ; Sequence 5, Application US/09835684  
 ; Patent No. US20020019337A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Wei, Zhong-Min  
 ; APPLICANT: Qiu, Dewen  
 ; APPLICANT: Remick, Dean  
 ; TITLE OF INVENTION: TREATMENT OF FRUITS OR VEGETABLES WITH HYPERSENSITIVE  
 ; TITLE OF INVENTION: RESPONSE ELICITOR TO CONTROL POSTHARVEST DISEASE OR  
 ; TITLE OF INVENTION: DESICCATION  
 ; FILE REFERENCE: 21829/71  
 ; CURRENT APPLICATION NUMBER: US/09/835,684  
 ; CURRENT FILING DATE: 2001-04-16  
 ; PRIOR APPLICATION NUMBER: 60/198,359  
 ; PRIOR FILING DATE: 2000-04-19  
 ; NUMBER OF SEQ ID NOS: 12  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 5  
 ; LENGTH: 447  
 ; TYPE: PRT  
 ; ORGANISM: Erwinia amylovora  
 US-09-835-684-5

Query Match 100.0%; Score 128; DB 9; Length 447;  
 Best Local Similarity 100.0%; Pred. No. 3e-10;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ITPDGGGGGQIGDNPLKAMKLIA 25  
 DB 116 ITPDGGGGGQIGDNPLKAMKLIA 140

RESULT 2  
 US-09-880-371-5  
 ; Sequence 5, Application US/09880371  
 ; Patent No. US20020059658A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Wei, Zhong-Min  
 ; APPLICANT: Derocher, Jay  
 ; TITLE OF INVENTION: METHODS OF IMPROVING THE EFFECTIVENESS OF TRANSGENIC  
 ; TITLE OF INVENTION: PLANTS  
 ; FILE REFERENCE: 21829/91  
 ; CURRENT APPLICATION NUMBER: US/09/880,371  
 ; CURRENT FILING DATE: 2001-06-13  
 ; PRIOR APPLICATION NUMBER: 60/211,585  
 ; PRIOR FILING DATE: 2000-06-15  
 ; NUMBER OF SEQ ID NOS: 16  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 5  
 ; LENGTH: 447  
 ; TYPE: PRT  
 ; ORGANISM: Erwinia amylovora  
 US-09-880-371-5

Query Match 100.0%; Score 128; DB 9; Length 447;  
 Best Local Similarity 100.0%; Pred. No. 3e-10;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ITPDGGGGGQIGDNPLKAMKLIA 25  
 DB 116 ITPDGGGGGQIGDNPLKAMKLIA 140

RESULT 3  
 US-09-879-248-6  
 ; Sequence 6, Application US/09879248  
 ; Patent No. US20020062500A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Fan, Hao  
 ; APPLICANT: Wei, Zhong-Min  
 ; TITLE OF INVENTION: HYPERSENSITIVE RESPONSE ELICITING DOMAINS AND USE  
 ; TITLE OF INVENTION: THEREOF  
 ; FILE REFERENCE: 21829/81  
 ; CURRENT APPLICATION NUMBER: US/09/879,248  
 ; CURRENT FILING DATE: 2001-06-12  
 ; PRIOR APPLICATION NUMBER: 60/212,211  
 ; PRIOR FILING DATE: 2000-06-16  
 ; NUMBER OF SEQ ID NOS: 18  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 6  
 ; LENGTH: 447  
 ; TYPE: PRT  
 ; ORGANISM: Erwinia amylovora  
 US-09-879-248-6

Query Match 100.0%; Score 128; DB 9; Length 447;  
 Best Local Similarity 100.0%; Pred. No. 3e-10;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ITPDGGGGGQIGDNPLKAMKLIA 25  
 DB 116 ITPDGGGGGQIGDNPLKAMKLIA 140

RESULT 4  
 US-10-010-390-5  
 ; Sequence 5, Application US/10010390  
 ; Publication No. US20030104979A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Wei, Zhong-Min  
 ; APPLICANT: Leon, Ernesto  
 ; APPLICANT: Oviedo, Agustín  
 ; TITLE OF INVENTION: METHODS OF INHIBITING DESICCATION OF CUTTINGS REMOVED  
 ; TITLE OF INVENTION: FROM ORNAMENTAL PLANTS  
 ; FILE REFERENCE: 21829/111  
 ; CURRENT APPLICATION NUMBER: US/10/010,390  
 ; CURRENT FILING DATE: 2001-11-05  
 ; PRIOR APPLICATION NUMBER: 60/248,169  
 ; PRIOR FILING DATE: 2000-11-13  
 ; NUMBER OF SEQ ID NOS: 14  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 5  
 ; LENGTH: 447  
 ; TYPE: PRT  
 ; ORGANISM: Erwinia amylovora  
 US-10-010-390-5

Query Match 100.0%; Score 128; DB 14; Length 447;  
 Best Local Similarity 100.0%; Pred. No. 3e-10;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ITPDGGGGGQIGDNPLKAMKLIA 25  
 DB 116 ITPDGGGGGQIGDNPLKAMKLIA 140



```
RESULT 5
US-10-441-736-6
; Sequence 6, Application US/10441736
; Publication No. US20040016029A1
; GENERAL INFORMATION:
; APPLICANT: Wei, Zhong-Min
; APPLICANT: Schading, Richard L.
; TITLE OF INVENTION: HYPERSENSITIVE RESPONSE ELICITOR-INDUCED STRESS
; TITLE OF INVENTION: RESISTANCE
; FILE REFERENCE: 21829/203 (EBC-003)
; CURRENT APPLICATION NUMBER: US/10/441,736
; CURRENT FILING DATE: 2003-05-20
; PRIOR APPLICATION NUMBER: 60/107,243
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: 09/431,614
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 6
; LENGTH: 447
; TYPE: PRT
; ORGANISM: Erwinia amylovora
US-10-441-736-6

Query Match      100.0%; Score 128; DB 15; Length 447;
Best Local Similarity 100.0%; Pred. No. 3e-10;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ITPDGGGGGGGIGDNPLLKMKLKLIA 25
DB      116 ITPDGGGGGGGIGDNPLLKMKLKLIA 140

RESULT 6
US-10-847-142-5
; Sequence 5, Application US/10847142
; Publication No. US20040285442A1
; GENERAL INFORMATION:
; APPLICANT: Wei, Zhong-Min
; APPLICANT: Qiu, Dewen
; APPLICANT: Remick, Dean
; TITLE OF INVENTION: TREATMENT OF FRUITS OR VEGETABLES WITH HYPERSENSITIVE
; TITLE OF INVENTION: RESPONSE ELICITOR TO CONTROL POSTHARVEST DISEASE OR
; TITLE OF INVENTION: DESICCATION
; FILE REFERENCE: 21829/197
; CURRENT APPLICATION NUMBER: US/10/847,142
; CURRENT FILING DATE: 2004-05-17
; PRIOR APPLICATION NUMBER: 60/198,359
; PRIOR FILING DATE: 2000-04-19
; PRIOR APPLICATION NUMBER: 09/835,684
; PRIOR FILING DATE: 2001-04-16
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 5
; LENGTH: 447
; TYPE: PRT
; ORGANISM: Erwinia amylovora
US-10-847-142-5

Query Match      100.0%; Score 128; DB 17; Length 447;
Best Local Similarity 100.0%; Pred. No. 3e-10;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ITPDGGGGGGGIGDNPLLKMKLKLIA 25
DB      116 ITPDGGGGGGGIGDNPLLKMKLKLIA 140

RESULT 7
US-10-109-670-10
; Sequence 10, Application US/10109670
; Publication No. US20030105283A1
; GENERAL INFORMATION:
; APPLICANT: TAKESAKO, KAZUTOH et al.
; TITLE OF INVENTION: ANTIGENIC PROTEIN ORIGINATING IN MALASSEZIA
; FILE REFERENCE: 1422-0523P
; CURRENT APPLICATION NUMBER: US/10/109,670
; CURRENT FILING DATE: 2002-04-01
; NUMBER OF SEQ ID NOS: 58
; SEQ ID NO 10
; LENGTH: 206
; TYPE: PRT
; ORGANISM: Malassezia furfur
US-10-109-670-10

Query Match      44.5%; Score 57; DB 14; Length 206;
Best Local Similarity 45.0%; Pred. No. 3.4;
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY      1 ITPDGGGGGGGIGDNPLLKAM 20
DB      86 LAPQSEGGQLNDGPKLQAI 105

RESULT 8
US-10-045-792-12
; Sequence 12, Application US/10045792
; Publication No. US20030003563A1
; GENERAL INFORMATION:
; APPLICANT: Vinkemeier, Uwe
; APPLICANT: Moarefi, Ismail
; APPLICANT: Darnell, Jr., James E.
; APPLICANT: Kurivan, John
; TITLE OF INVENTION: A CRYSTAL OF THE N-TERMINAL DOMAIN OF A
; TITLE OF INVENTION: STAT PROTEIN AND METHODS OF USE THEREOF
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue, 4th Floor
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/045,792
; FILING DATE: 19-Oct-2001
; CLASSIFICATION: <unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/012,710
; FILING DATE: <unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-194
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 837 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-045-792-12

Query Match      42.2%; Score 54; DB 14; Length 837;
Best Local Similarity 64.3%; Pred. No. 42;
```

```
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 5 GGGGGQIGDNPLK 18
Db 802 GEGGSLGSOPLK 815

RESULT 9
US-09-845-335-3
; Sequence 3, Application US/09845335
; Patent No. US20020058266A1
; GENERAL INFORMATION:
; APPLICANT: CLOUGH, BARBARA
; APPLICANT: PREISER, PETER
; APPLICANT: WILSON, ROBERT
; TITLE OF INVENTION: AN EP-TU PROTEIN ENCODED ON THE PLASTID DNA OF THE
; TITLE OF INVENTION: MALARIA PARASITE AND PROTEIN SYNTHESIS INHIBITORS
; TITLE OF INVENTION: EFFECTIVE AS ANTI-MALARIAL COMPOUNDS
; FILE REFERENCE: N68837B GCW EIC DP
; CURRENT APPLICATION NUMBER: US/09/845,335
; CURRENT FILING DATE: 2001-12-26
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/140,466
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-08-26
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 409
; TYPE: PRT
; ORGANISM: Anacystis nidulans
US-09-845-335-3

Query Match 39.8%; Score 51; DB 9; Length 409;
Best Local Similarity 57.1%; Pred. No. 53;
Matches 12; Conservative 3; Mismatches 4; Indels 2; Gaps 1;

QY 6 QGG--GQIGDNPLKMLKLI 24
Db 183 QGGASGQGDNPWVDKILKM 203

RESULT 10
US-10-424-599-214915
; Sequence 214915, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 214915
; LENGTH: 78
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_36094C.1.pep
US-10-424-599-214915

Query Match 39.1%; Score 50; DB 15; Length 78;
Best Local Similarity 61.1%; Pred. No. 13;
Matches 11; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2 TPDGCGGQIGDNPLKA 19
Db 40 TPKGGGGGPRGNPNKA 57

RESULT 11
US-09-879-248-6_copy_116_140.rapb
```

```
US-10-437-963-149121
; Sequence 149121, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 149121
; LENGTH: 177
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_49486C.1.pep
US-10-437-963-149121

Query Match 39.1%; Score 50; DB 16; Length 177;
Best Local Similarity 47.6%; Pred. No. 31;
Matches 10; Conservative 4; Mismatches 7; Indels 0; Gaps 0;
```

```
QY 5 GGGGGQIGDNPLKMLKLI 25
Db 19 GVGCGRLGQGPSNEASRLAA 39
```

```
RESULT 12
US-10-425-115-231795
; Sequence 231795, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 231795
; LENGTH: 149
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(149)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_142992C.1.pep
US-10-425-115-231795

Query Match 38.3%; Score 49; DB 17; Length 149;
Best Local Similarity 66.7%; Pred. No. 36;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 3 PDGCGGQIGDN 14
Db 65 PDGEGGQTD 76
```

```
RESULT 13
US-10-369-493-3880
```

```

, , PRIOR APPLICATION NUMBER: US01/013202
, , PRIOR FILING DATE: 2001-01-17
, , PRIOR APPLICATION NUMBER: 09/764,853
, , PRIOR FILING DATE: 2001-01-17
, , PRIOR APPLICATION NUMBER: US01/01349
, , PRIOR FILING DATE: 2001-01-17
, , PRIOR APPLICATION NUMBER: 09/764,902
, , PRIOR FILING DATE: 2001-01-17
, , PRIOR APPLICATION NUMBER: US01/01239
, , PRIOR FILING DATE: 2001-01-17
, , PRIOR APPLICATION NUMBER: 09/764,870
, , PRIOR FILING DATE: 2001-01-17
, , PRIOR APPLICATION NUMBER: US01/01348
, , PRIOR FILING DATE: 2001-01-17
, , PRIOR APPLICATION NUMBER: 09/764,882
, , PRIOR FILING DATE: 2001-01-17
, , PRIOR APPLICATION NUMBER: US01/01347
, , PRIOR FILING DATE: 2001-01-17
, , PRIOR APPLICATION NUMBER: 09/764,896
, , PRIOR FILING DATE: 2001-01-17
, , PRIOR APPLICATION NUMBER: US01/01307
, , PRIOR FILING DATE: 2001-01-17
, , PRIOR APPLICATION NUMBER: 09/764,864
, , PRIOR FILING DATE: 2001-01-17
, , PRIOR APPLICATION NUMBER: US01/01336
, , PRIOR FILING DATE: 2001-01-17
, , PRIOR APPLICATION NUMBER: 09/764,868
, , PRIOR FILING DATE: 2001-01-17
, , PRIOR APPLICATION NUMBER: US01/01312
, , PRIOR FILING DATE: 2001-01-17
, , PRIOR APPLICATION NUMBER: 60/179,065
, , PRIOR FILING DATE: 2000-01-31
, , PRIOR APPLICATION NUMBER: 60/180,628
, , PRIOR FILING DATE: 2000-02-04
, , PRIOR APPLICATION NUMBER: 60/209,467
, , PRIOR FILING DATE: 2000-06-07
, , NUMBER OF SEQ ID NOS: 167
, , SOFTWARE: Patentin Ver. 2.0
, , SEQ ID NO 100
, , LENGTH: 365
, , TYPE: PRT
, , ORGANISM: Homo sapiens
, , US-09-908-711-100

Query Match 37.9%; Score
Best Local Similarity 54.5%; Pred.
Matches 12; Conservative 4; Mismatches 15

QY 2 TPDQGGGGIGDGNPLIKAMKL 23
||| : ||| ||||| : :
Db 284 TPDKGQTQI-ENPLIKNLRV 304

RESULT 15
US-10-269-806-163
, , Sequence 163, Application US/10269806
, , Publication No. US20030176352A1
, , GENERAL INFORMATION:
, , APPLICANT: Min, Housung
, , APPLICANT: Sitney, Karen
, , APPLICANT: Hartley, Cynthia
, , TITLE OF INVENTION: Peptides and Related Compounds
, , FILE REFERENCE: A-750
, , CURRENT APPLICATION NUMBER: US/10/269
, , CURRENT FILING DATE: 2002-10-10
, , NUMBER OF SEQ ID NOS: 199
, , SOFTWARE: Patentin version 3.1
, , SEQ ID NO 163
, , LENGTH: 43

```

```
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthesized Peptide Sequence
; NAME/KEY: misc feature
; LOCATION: (1)..(1)
; OTHER INFORMATION: At position 1, Pc at N-terminus
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (44)..(44)
; OTHER INFORMATION: At position 44, Pc at C-terminus
US-10-269-806-163
```

```
Query Match 37.5%; Score 48; DB 14; Length 43;
Best Local Similarity 50.0%; Pred. No. 13;
Matches 9; Conservative 3; Mismatches 6; Indels 0; Gaps 0;
```

```
QY 5 GGGGGQIGDNPLKAMLK 22
Db 21 GRGGGALRDGPTLKQWLE 38
```

```
RESULT 16
US-10-437-963-124136
; Sequence 124136, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 124136
; LENGTH: 111
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(111)
; OTHER INFORMATION: unsure at all xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_26904C.1.pap
US-10-437-963-124136
```

```
Query Match 37.5%; Score 48; DB 16; Length 111;
Best Local Similarity 55.6%; Pred. No. 36;
Matches 10; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
```

```
QY 5 GGGGGQIGDNPLKAMLK 22
Db 4 GPGGGKIGKPLRVGLLK 21
```

```
RESULT 17
US-10-032-585-7297
; Sequence 7297, Application US/10032585
; Publication No. US20030180953A1
; GENERAL INFORMATION:
; APPLICANT: Terry, Roemer D.
; APPLICANT: Bo, Jiang
; APPLICANT: Charles, Boone
; APPLICANT: Howard, Bussey
; TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
```

```
; FILE REFERENCE: 10182-005-999
; CURRENT APPLICATION NUMBER: US/10/032,585
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 8000
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7297
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Candida albicans
US-10-032-585-7297
```

```
Query Match 37.5%; Score 48; DB 14; Length 238;
Best Local Similarity 43.5%; Pred. No. 82;
Matches 10; Conservative 4; Mismatches 9; Indels 0; Gaps 0;
```

```
QY 2 TPDGGGGQIGDNPLKAMLKLI 24
Db 50 TTDGNGGGGIGIVPTLQNVATV 72
```

```
RESULT 18
US-10-425-115-301902
; Sequence 301902, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 301902
; LENGTH: 207
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_3840C.1.pap
US-10-425-115-301902
```

```
Query Match 37.1%; Score 47.5; DB 17; Length 207;
Best Local Similarity 64.7%; Pred. No. 84;
Matches 11; Conservative 2; Mismatches 3; Indels 1; Gaps 1;
```

```
QY 5 GGGGGQIGDNPLKAMLK 21
Db 77 GAGGGQDGDGDE-LMRAML 92
```

```
RESULT 19
US-10-425-115-347563
; Sequence 347563, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 347563
; LENGTH: 81
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_8013C.1.pap
```

US-10-425-115-347563

Query Match 36.7%; Score 47; DB 17; Length 81;  
 Best Local Similarity 40.0%; Pred. No. 36;  
 Matches 8; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 1 ITPDGGGGQIGDNPILKAM 20  
 Db 17 VSPGALGKGQIGSRPVLNGL 36

RESULT 20

US-10-424-599-216781  
 ; Sequence 216781, Application US/10424599  
 ; Publication No. US20040031072A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: La Rosa Thomas J  
 ; APPLICANT: Kovalic David K  
 ; APPLICANT: Zhou Yihua  
 ; APPLICANT: Cao Yongwei  
 ; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
 ; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
 ; FILE REFERENCE: 38-21(53223)B  
 ; CURRENT APPLICATION NUMBER: US/10/424,599  
 ; CURRENT FILING DATE: 2003-04-28  
 ; NUMBER OF SEQ ID NOS: 285684  
 ; SEQ ID NO 216781  
 ; LENGTH: 102  
 ; TYPE: PRT  
 ; ORGANISM: Glycine max  
 ; FEATURE:  
 ; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_37782C.1.pep  
 US-10-424-599-216781

Query Match 36.7%; Score 47; DB 15; Length 102;  
 Best Local Similarity 50.0%; Pred. No. 47;  
 Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 5 GGGGGQIGDNPILKAMLK 22  
 Db 24 GGGGGQLGESPWGEAFLK 41

Search completed: January 25, 2005, 12:18:59  
 Job time : 51.6346 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2005, 12:08:06 ; Search time 11.0577 Seconds  
(without alignments)  
217.534 Million cell updates/sec

Title: US-09-879-248-6\_COPY\_116\_140  
Perfect score: 128  
Sequence: 1 ITPDGGGQIGDNPLLKWLKIIA 25

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 100 summaries

Database : PIR 79:.\*  
1: piri:.\*  
2: pir2:.\*  
3: pir3:.\*  
4: pir4:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	128	100.0	447	2 T18447	HrpW protein - Erw
2	54	42.2	837	2 I57557	DNA-Binding Protei
3	52	40.6	135	2 B49218	hemagglutinin homo
4	52	40.6	702	2 S48753	major surface prot
5	51	39.8	409	2 S04430	translation elonga
6	49.5	38.7	480	2 E75433	hypothetical prote
7	49	38.3	428	2 G71177	hypothetical prote
8	49	38.3	436	2 T46107	hypothetical prote
9	49	38.3	598	2 S66669	potassium channel
10	49	38.3	613	2 A56031	potassium channel
11	48	37.5	225	2 AF2144	hypothetical prote
12	47	36.7	435	2 T20152	hypothetical prote
13	47	36.7	1343	1 H64073	DNA-directed RNA p
14	46.5	36.3	652	2 D82317	iron-regulated out
15	46.5	36.3	857	2 H70411	alanine-tRNA ligas
16	46	35.9	435	2 S84674	probable AAA-type
17	46	35.9	605	2 S01066	regulatory protein
18	46	35.9	1134	2 D75014	hypothetical prote
19	45.5	35.5	161	2 T42006	troponin C - Caeno
20	45.5	35.5	694	2 S09794	hypothetical prote
21	45	35.2	206	2 JC2574	hypothetical 22K p
22	45	35.2	253	2 F86184	hypothetical prote
23	45	35.2	283	2 T19933	hypothetical prote
24	45	35.2	302	2 A84263	hypothetical prote
25	45	35.2	373	2 A56392	beta-galactoside a
26	45	35.2	405	2 D75563	hypothetical prote
27	45	35.2	405	2 AD2784	molybdopterin bios
28	45	35.2	509	2 D83651	hypothetical prote
29	45	35.2	532	2 A90037	hypothetical prote

30	45	35.2	576	2 T17842	hypothetical prote
31	45	35.2	664	1 VEHULA	lamin A - human
32	45	35.2	665	2 S27267	lamin A - rat
33	45	35.2	665	2 S28182	lamin A - mouse
34	45	35.2	1547	2 T13847	sno protein - frui
35	44	34.4	258	2 G72313	hypothetical prote
36	44	34.4	394	2 D69370	probable acyl-CoA
37	44	34.4	412	2 T32890	steroid/thyroid/re
38	44	34.4	501	2 A11031	probable membrane
39	44	34.4	535	2 S44827	F54P2.2 protein -
40	44	34.4	790	2 E48327	COI intron A prote
41	44	34.4	875	2 AE0401	alanine-tRNA ligas
42	44	34.4	890	2 P75289	alanyl-tRNA synthe
43	43.5	34.0	415	2 S60078	Runt domain contai
44	43.5	34.0	415	2 B55563	AML2a protein - hu
45	43.5	34.0	754	2 S41391	gelsolin - America
46	43.5	34.0	754	2 S53373	gelsolin - America
47	43.5	34.0	962	2 AG2444	hypothetical prote
48	43	33.6	182	2 T18709	hypothetical prote
49	43	33.6	254	2 B72374	conserved hypotet
50	43	33.6	290	1 H70057	agmatinase homolog
51	43	33.6	328	2 E81452	probable fatty aci
52	43	33.6	341	2 AC3638	tetraacyldisacchar
53	43	33.6	410	2 G69274	probable acyl-CoA
54	43	33.6	414	1 JN0826	probable beta-keto
55	43	33.6	419	1 JC1211	beta-ketoacyl synt
56	43	33.6	437	2 F83854	hypothetical prote
57	43	33.6	454	2 AE1970	hypothetical prote
58	43	33.6	584	2 S62198	formylmethanofuran
59	43	33.6	597	2 S51212	BAK5 protein - bov
60	43	33.6	615	1 S08977	nif-specific regul
61	43	33.6	740	2 T03975	hypothetical prote
62	43	33.6	877	2 S76394	hypothetical prote
63	43	33.6	892	2 D69393	large helicase-rel
64	43	33.6	900	2 H90176	alanyl-tRNA synthe
65	42.5	33.2	239	2 S07377	ribosomal protein
66	42.5	33.2	253	1 R52PD4	ribosomal protein
67	42.5	33.2	265	2 H89834	ferrichrome transp
68	42.5	33.2	291	2 H71467	succinate-CoA liga
69	42.5	33.2	373	2 E72216	alanine-tRNA ligas
70	42.5	33.2	413	2 AC3614	tdp-4-dehydroham
71	42.5	33.2	652	2 S25265	outer membrane pro
72	42.5	33.2	717	2 S77190	hypothetical prote
73	42.5	33.2	860	2 A96717	unknown protein, 4
74	42	32.8	67	2 B35063	hypothetical prote
75	42	32.8	122	1 B69067	conserved hypotet
76	42	32.8	180	2 C86751	hypothetical prote
77	42	32.8	185	2 AD0834	hypothetical prote
78	42	32.8	233	1 DSBYN	superoxide dismuta
79	42	32.8	242	2 G81058	conserved hypotet
80	42	32.8	247	2 AG0250	conserved hypotet
81	42	32.8	260	1 A32515	acetolactate decar
82	42	32.8	266	2 S56520	hypothetical 29.1K
83	42	32.8	268	2 B64066	probable ABC trans
84	42	32.8	317	2 A64343	hypothetical prote
85	42	32.8	335	2 G82090	transcription regu
86	42	32.8	350	2 I55214	salivary proline-r
87	42	32.8	366	2 S77203	hypothetical prote
88	42	32.8	374	2 T36363	probable DNA-bindi
89	42	32.8	396	2 T01049	hypothetical prote
90	42	32.8	408	2 G87646	molybdopterin bios
91	42	32.8	409	2 H70303	ubiquinol-cytochro
92	42	32.8	459	2 C64679	phosphomannomutase
93	42	32.8	459	2 T41840	phosphohexosemutas
94	42	32.8	462	2 T42679	hypothetical prote
95	42	32.8	492	2 S28007	probable ATP-bindi
96	42	32.8	508	2 T09046	proline-rich prote
97	42	32.8	529	2 S12787	potassium channel
98	42	32.8	541	2 JX0112	signal recognition
99	42	32.8	546	2 T37665	probable t-complex
100	42	32.8	664	2 A56695	notch2 protein hom

## ALIGNMENTS

## RESULT 1

T18447  
HrpW protein - Erwinia amylovora  
C:Species: Erwinia amylovora  
C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: T18447  
R:Gaudriault, S.  
submitted to the EMBL Data Library, May 1998  
A:Reference number: Z18936  
A:Accession: T18447  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-447 <GAU>  
A:Cross-references: UNIPROT:O54508; EMBL:Y13831; NID:e1293450; PID:e1293451; PIDN:CAA741  
A:Experimental source: strain CFBP1430; specific host Pommoideae  
C:Genetics:  
A:Note: hrpW

Query Match 100.0%; Score 128; DB 2; Length 447;  
Best Local Similarity 100.0%; Pred. No. 1e-11;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ITPDGGGGGQIGDNP LKAMLKLI A 25  
Db 116 ITPDGGGGGQIGDNP LKAMLKLI A 140

## RESULT 2

I57557  
DNA-Binding protein and transcription factor - mouse  
C:Species: Mus musculus (house mouse)  
C>Date: 29-May-1998 #sequence\_revision 29-May-1998 #text\_change 09-Jul-2004  
C:Accession: I57557  
R:Quelle, F.W.; Shimoda, K.; Thierfelder, W.; Fischer, C.L.; Kim, A.; Ruben, S.M.; Cleve  
Mol. Cell. Biol. 15, 3336-3343, 1995  
A:Title: Cloning of murine Stat6 and human Stat6, Stat proteins that are tyrosine phosph  
A:Reference number: I57557; MUID:95280934; PMID:7760829  
A:Accession: I57557  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-837 <RES>  
A:Cross-references: UNIPROT:P52633; GB:L47650; NID:g1008876; PIDN:AAA79006.1; PID:g10088  
C:Genetics:  
A:Gene: STAT6  
C:Superfamily: human signal transducer and transcription activator STAT5A

Query Match 42.2%; Score 54; DB 2; Length 837;  
Best Local Similarity 64.3%; Pred. No. 5;  
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 5 GGGGGGQIGDNP LK 18  
Db 802 GGGGGGQIGDNP LK 815

## RESULT 3

B49218  
hemagglutinin homolog pmGA1.3 - Mycoplasma gallisepticum (fragment)  
C:Species: Mycoplasma gallisepticum  
C>Date: 21-Jan-1994 #sequence\_revision 18-Nov-1994 #text\_change 09-Jul-2004  
C:Accession: B49218  
R:Markham, P.F.; Glew, M.D.; Whithear, K.G.; Walker, I.D.  
Infect. Immun. 61, 903-909, 1993  
A:Title: Molecular cloning of a member of the gene family that encodes pmGA, a hemagglut  
A:Reference number: A49218; MUID:93162830; PMID:8432610  
A:Accession: B49218  
A:Status: preliminary  
A:Molecule type: DNA; protein  
A:Residues: 1-135 <WAR>  
A:Cross-references: UNIPROT:Q53304; GB:S55216; NID:g265625; PIDN:AAB25398.1; PID:g265627

A:Experimental source: S6  
A:Note: sequence extracted from NCBI backbone (NCBIN:125182, NCBIP:125184)  
C:Genetics:  
A:Genetic code: SGC3

Query Match 40.6%; Score 52; DB 2; Length 135;  
Best Local Similarity 64.3%; Pred. No. 1.3;  
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 TPDGGGGGQIGDNP 15  
Db 57 TNPNGGGGMMGDNP 70

## RESULT 4

S48753  
major surface protein (clone pmGA1.3) precursor - Mycoplasma gallisepticum  
C:Species: Mycoplasma gallisepticum  
C>Date: 01-Aug-1995 #sequence\_revision 03-Nov-1995 #text\_change 09-Jul-2004  
C:Accession: S48753  
R:Markham, P.F.; Glew, M.D.; Sykes, J.E.; Bowden, T.R.; Pollocks, T.D.; Browning, G.F.; V  
FEBS Lett. 352, 347-352, 1994  
A:Title: The organisation of the multigene family which encodes the major cell surface pr  
A:Reference number: S48751; MUID:95010739; PMID:7925999  
A:Accession: S48753  
A:Status: preliminary; nucleic acid sequence not shown  
A:Molecule type: DNA  
A:Residues: 1-702 <WAR>  
A:Cross-references: UNIPROT:Q49498; EMBL:L28424; NID:g535687; PIDN:AAA62417.1; PID:g53568  
C:Genetics:  
A:Genetic code: SGC3  
A:Start codon: GTG

Query Match 40.6%; Score 52; DB 2; Length 702;  
Best Local Similarity 64.3%; Pred. No. 8.3;  
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 TPDGGGGGQIGDNP 15  
Db 57 TNPNGGGGMMGDNP 70

## RESULT 5

S04430  
translation elongation factor EF-Tu.A - Synechococcus sp.  
C:Species: Synechococcus sp.  
C>Date: 28-Feb-1990 #sequence\_revision 28-Feb-1990 #text\_change 09-Jul-2004  
C:Accession: S04430; G60663  
R:Meng, B.Y.; Shinozaki, K.; Sugiura, M.  
Mol. Gen. Genet. 216, 25-30, 1989  
A:Title: Genes for the ribosomal proteins S12 and S7 and elongation factors EF-G and EF-T  
A:Reference number: S04426; MUID:89281486; PMID:2499762  
A:Accession: S04430  
A:Status: not compared with conceptual translation  
A:Molecule type: DNA  
A:Residues: 1-409 <MEN>  
A:Cross-references: UNIPROT:P18668  
A:Note: the source is designated as Anacystis nidulans  
R:Ludwig, W.; Weizenegger, M.; Betzl, D.; Leidel, E.; Lenz, T.; Ludvigsen, A.; Moellenhof  
Arch. Microbiol. 153, 241-247, 1990  
A:Title: Complete nucleotide sequences of seven eubacterial genes coding for the elongat  
A:Reference number: A60663; MUID:90240875; PMID:2110445  
A:Accession: G60663  
A:Status: not compared with conceptual translation  
A:Molecule type: DNA  
A:Residues: 1-324,'D',326-409 <LUD>  
A:Cross-references: UNIPROT:Q53304; GB:S55216; NID:g265625; PIDN:AAB25398.1; PID:g265627

A:Note: the source is designated as Anacystis nidulans  
R:Ludwig, W.; Weizenegger, M.; Betzl, D.; Leidel, E.; Lenz, T.; Ludvigsen, A.; Moellenhof  
Arch. Microbiol. 153, 241-247, 1990  
A:Title: Complete nucleotide sequences of seven eubacterial genes coding for the elongat  
A:Reference number: A60663; MUID:90240875; PMID:2110445  
A:Accession: G60663  
A:Status: not compared with conceptual translation  
A:Molecule type: DNA  
A:Residues: 1-324,'D',326-409 <LUD>  
A:Cross-references: UNIPROT:Q53304; GB:S55216; NID:g265625; PIDN:AAB25398.1; PID:g265627

A:Note: the source is designated as Anacystis nidulans  
R:Ludwig, W.; Weizenegger, M.; Betzl, D.; Leidel, E.; Lenz, T.; Ludvigsen, A.; Moellenhof  
Arch. Microbiol. 153, 241-247, 1990  
A:Title: Complete nucleotide sequences of seven eubacterial genes coding for the elongat  
A:Reference number: A60663; MUID:90240875; PMID:2110445  
A:Accession: G60663  
A:Status: not compared with conceptual translation  
A:Molecule type: DNA  
A:Residues: 1-324,'D',326-409 <LUD>  
A:Cross-references: UNIPROT:Q53304; GB:S55216; NID:g265625; PIDN:AAB25398.1; PID:g265627



F;174-176/Region: GTP-binding SAK/L motif

F;25,26,62,136,137,139,174/Binding site: Mg-GTP (Lys, Thr, Thr, Asn, Lys, Asp, Ser) #sta

Query Match 39.8%; Score 51; DB 2; Length 409;  
Best Local Similarity 57.1%; Pred. No. 6.5;  
Matches 12; Conservative 3; Mismatches 4; Indels 2; Gaps 1;

QY 6 QGG--GGIGNPPLKAMKLKI 24  
||| ||||| : |||:  
DB 183 QGGASGGQGNPWVDKILKLM 203

## RESULT 6

E75433  
hypothetical protein - Deinococcus radiodurans (strain R1)  
C;Species: Deinococcus radiodurans  
C;Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
C;Accession: E75433  
R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;  
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.  
S.; Shen, M.; Vamathevan, J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma  
Science 286, 1571-1577, 1999  
A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.  
A;Reference number: A75250; MUID:20036896; PMID:10567266  
A;Accession: E75433  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-480 <WHI>  
A;Cross-references: UNIPROT:Q9RV87; GB:AE001963; GB:AE000513; NID:96458869; PIDN:AAF1071  
A;Experimental source: strain R1  
C;Genetics:  
A;Gene: DR142  
A;Map position: 1  
C;Superfamily: Deinococcus radiodurans hypothetical protein DR1142

Query Match 38.7%; Score 49.5; DB 2; Length 480;  
Best Local Similarity 60.0%; Pred. No. 13;  
Matches 12; Conservative 2; Mismatches 5; Indels 1; Gaps 1;

QY 5 GGGGGGIGD-NPLLKAMKL 23  
||||| : ||||| :  
DB 157 GGGGALGDFAPILLAFRL 176

## RESULT 7

G71177  
hypothetical protein PH1701 - Pyrococcus horikoshii  
C;Species: Pyrococcus horikoshii  
C;Date: 14-Aug-1998 #sequence\_revision 14-Aug-1998 #text\_change 09-Jul-2004  
C;Accession: G71177  
R;Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekir  
M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi  
DNA Res. 5, 55-76, 1998  
A;Title: Complete sequence and gene organization of the genome of a hyper-thermophilic a  
A;Reference number: A71000; MUID:98344137; PMID:9679194  
A;Accession: G71177  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-428 <KAW>  
A;Cross-references: UNIPROT:O59360; GB:AP000007; NID:g3236134; PIDN:BAA30814.1; PID:g323  
A;Experimental source: strain OT3  
C;Note: this accession replaces an interim accession for a sequence replaced by GenBank  
C;Genetics:  
A;Gene: PH1701

Query Match 38.3%; Score 49; DB 2; Length 428;  
Best Local Similarity 62.5%; Pred. No. 14;  
Matches 10; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 8 GGGIGNPPLKAMKL 23  
||| ||||| : |||:  
DB 286 GAQSGONAMUKAMHRL 301

## RESULT 8

T46107  
hypothetical protein T25B15.140 - Arabidopsis thaliana  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 04-Feb-2000 #sequence\_revision 04-Feb-2000 #text\_change 09-Jul-2004  
C;Accession: T46107  
R;Alcaraz, J.P.; Clabault, G.; Cottet, A.; Mache, R.; Mewes, H.W.; Lemcke, K.; Mayer, K.  
submitted to the Protein Sequence Database, January 2000  
A;Reference number: Z23021  
A;Accession: T46107  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-436 <ALC>  
A;Cross-references: UNIPROT:Q9FT45; EMBL:AL132972  
A;Experimental source: cultivar Columbia; BAC clone T25B15  
C;Genetics:  
A;Map position: 3  
A;Introns: 418/1  
A;Note: T25B15.140

Query Match 38.3%; Score 49; DB 2; Length 436;  
Best Local Similarity 52.6%; Pred. No. 14;  
Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 3 PDGGGGGQIGDNPPLKAML 21  
||| ||||| : |||:  
DB 22 PDKPGSGQINSNSVLVALL 40

## RESULT 9

S66669  
potassium channel (Kv1.5) - rabbit  
C;Species: Oryctolagus cuniculus (domestic rabbit)  
C;Date: 28-Oct-1996 #sequence\_revision 13-Mar-1997 #text\_change 09-Jul-2004  
C;Accession: S66669  
R;Sasaki, Y.; Ishii, K.; Nunoki, K.; Yamagishi, T.; Taira, N.  
FEBS Lett. 372, 20-24, 1995  
A;Title: The voltage-dependent K(+) channel (Kv1.5) cloned from rabbit heart and facilitat  
A;Reference number: S66669; MUID:96032538; PMID:7556635  
A;Accession: S66669  
A;Status: preliminary; nucleic acid sequence not shown  
A;Molecule type: mRNA  
A;Residues: 1-598 <SAS>  
A;Cross-references: UNIPROT:P50638; EMBL:D45025; NID:g1060972; PIDN:BAA08082.1; PID:d1006  
C;Superfamily: potassium channel protein drkl

Query Match 38.3%; Score 49; DB 2; Length 598;  
Best Local Similarity 45.5%; Pred. No. 20;  
Matches 10; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 3 PDGGGGGQIGDNPPLKAMKLKI 24  
||| ||||| : |||:  
DB 366 PDGGGGGQGGQQAAMSLILRVI 387

## RESULT 10

A56031  
potassium channel KCNA5 - human  
N;Alternate names: potassium channel HK2; potassium channel PCN1; shaker-related potassi  
C;Species: Homo sapiens (man)  
C;Date: 05-Apr-1995 #sequence\_revision 05-Apr-1995 #text\_change 09-Jul-2004  
C;Accession: A56031; A38556; B39922; A38074  
R;Philipson, L.H.; LaMendola, J.; Bell, G.I.; Steiner, D.F.  
submitted to GenBank, September 1990  
A;Reference number: A56031  
A;Accession: A56031  
A;Molecule type: mRNA  
A;Residues: 1-613 <PHI>  
A;Cross-references: UNIPROT:P22460; GB:M55513; NID:g189653; PIDN:AAA36422.1; PID:g189654  
R;Philipson, L.H.; Hice, R.E.; Schaefer, K.; LaMendola, J.; Bell, G.I.; Neelson, D.J.; St  
Proc. Natl. Acad. Sci. U.S.A. 88, 53-57, 1991  
A;Title: Sequence and functional expression in Xenopus oocytes of a human insulinoma and

A:Reference number: A38556; MUID:91095456; PMID:1986382  
A:Accession: A38556  
A>Status: nucleic acid sequence not shown  
A:Molecule type: DNA  
A:Residues: 1-56, 'G', 58-137, 'L', 139-213, 'R', 215-227, 'P', 229-545, 'QG', 546-613 <PH2>  
A:Cross-references: GB:M55513  
R:Tamkun, M.M.; Knoch, K.B.; Walbridge, J.A.; Kroemer, H.; Roden, D.M.; Glover, D.M.  
FASEB J. 5, 331-337, 1991  
A>Title: Molecular cloning and characterization of two voltage-gated K(+) channel cDNAs  
A:Reference number: A39922; MUID:91160866; PMID:2001794  
A:Accession: B39922  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-54, 56-137, 'L', 139-186, 'G', 189-213, 'R', 215-227, 'P', 229-297, 'PTQRGH', 309-558  
A:Cross-references: GB:M60451  
R:Curran, M.E.; Landes, G.W.; Keating, M.T.  
Genomics 12, 729-737, 1992  
A>Title: Molecular cloning, characterization, and genomic localization of a human potas  
A:Reference number: A38074; MUID:92241872; PMID:1348297  
A:Accession: A38074  
A:Molecule type: DNA  
A:Residues: 1-137, 'L', 139-153, 'R', 155-213, 'R', 215-227, 'P', 229-281, 'V', 283-578, 'QLPPREV'  
A:Cross-references: GB:M83254; MID:g190202; PIDN:AAA60146.1; PID:g190203  
A:Experimental source: heart  
A>Note: sequence extracted from NCBI backbone (NCBIN:98573, NCBIP:98577)  
C:Genetics:  
A:Gene: GDB:KCNA5  
A:Cross-references: GDB:127904; OMIM:176267  
A:Map position: 12p13.33-12p13.31  
C:Superfamily: potassium channel protein drkl  
C:Keywords: glycoprotein; phosphoprotein; potassium channel; transmembrane protein; volt  
F:125,190/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F:557/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 38.3%; Score 49; DB 2; Length 613;  
Best Local Similarity 45.5%; Fred. No. 21;  
Matches 10; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 3 PDGGGGGGIGNDPLLLKMLKLI 24  
| | | | | : | : | : |  
DB 381 PGCGGGGGGQQQAMSLAIRVI 402

RESULT 11  
AF2140  
Hypothetical protein all2677 [imported] - Nostoc sp. (strain PCC 7120)  
C:Species: Nostoc sp. PCC 7120  
A>Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
C>Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Jul-2004  
C:Accession: AF2140  
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,  
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.  
DNA Res. 8, 205-213, 2001  
A>Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana  
A:Reference number: AB1807; MUID:21595285; PMID:11759840  
A:Accession: AF2140  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-225 <KUR>  
A:Cross-references: UNIPROT:Q8YTN7; GB:BA000019; PIDN:BAB74376.1; PID:g17131770; GSPDB:C  
A:Experimental source: strain PCC 7120  
C:Genetics:  
A:Gene: all2677

```
Query Match      37.5%; Score 48; DB 2; Length 225;
Best Local Similarity 43.5%; Pred. No. 9.7;
Matches 10; Conservative 4; Mismatches 9; Indels 0; Gaps 0;
```

Qy 3 PDQGCGGQTGDNPLLKAMLKLIA 25  
| : | | | | : | : |  
Db 137 PEAQTGGIGCTPLTQPHPQLVA 159

RESULT 12

T20152

hypothetical protein C52E4.4 - *Caenorhabditis elegans*

C:Species: *Caenorhabditis elegans*

C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004

C:Accession: T20152

R:Holt, R.

submitted to the EMBL Data Library, August 1996

A:Reference number: Z19230

A:Accession: T20152

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-435 <WIL>

A:Cross-references: UNIPROT:Q18787; EMBL:Z78012; PIDN:CA801414.1; GSPDB:GN01

A:Experimental source: clone C52E4

C:Genetics:

A:Gene: CESP:C52E4.4

A:Map position: 5

A:Introns: 26/1.60/3; 137/3; 318/3

C:Success:

Query Match	36.7%;	Score 47;	DB 2;	Length 435;
Best Local Similarity	52.4%;	Pred. No. 29;		
Matches 11;	Conservative 3;	Mismatches 3;	Indels 4;	Gaps 1;

QY 4 DQGGGQIGDNPLLKAMLKLI 24  
||||| : : : :  
Db 289 DQGQ---GDNEVORTMLELI 309

RESULT 13  
H64073

DNA-directed RNA polymerase (EC 2.7.7.6) beta chain - Haemophilus influenzae (strain Rd)  
C:Species: Haemophilus influenzae  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 09-Jul-2004  
C:Accession: H64073  
R:Rietschmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J. D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M. Science 269, 496-512, 1995  
A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter, A.  
A>Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.  
A:Reference number: A64000; MUID:95350630; PMID:7542800  
A:Accession: H64073  
A>Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-1343 <TIGR>  
A:Cross-references: UNIPROT:P43738; GB:U32733; GB:I42023; NID:g1573493; PIDN:AAC22173.1;  
C:Superfamily: DNA-directed RNA polymerase beta chain  
C:Keywords: nucleotidyltransferase; transcription

Query Match	36.7%	Score 47;	DB 1;	Length 1343;
Best Local Similarity	36.4%;	Pred. NO. 1e+02;		
Matches 8;	Conservative	6;	Mismatches	8; Indels 0; Gaps 0;

QY	3	PDQGGQIGDNPLLKAMKLI	24
		:   :   :   :	
db	412	PEGEGAGILSNEDIIAVMRKLI	433

RESULT 14  
D82317

Iron-regulated outer membrane virulence protein, TonB receptor family VC0475 [imported]  
 Species: *Vibrio cholerae*  
 Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 09-Jul-2004  
 Accession: D82317  
 Heideberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;  
 Richardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers,  
 R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
 Nature 406, 477-483, 2000  
 Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.  
 Reference number: A82035; MUID:20406833; PMID:10952301  
 Accession: D82317

Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 4 DQGGGGQIGDNP 15  
 ||:||||: ||:|  
 Db 98 DGGGGKGDGDP 109

RESULT 17  
 S01066  
 regulatory protein nifa - Bradyrhizobium japonicum  
 C:Species: Bradyrhizobium japonicum  
 C:Date: 31-Dec-1988 #sequence\_revision 31-Dec-1988 #text\_change 28-Apr-2003  
 C:Accession: S01066  
 R:Thoeny, B.; Fischer, H.M.; Anthamatten, D.; Bruderer, T.; Hennecke, H.  
 Nucleic Acids Res. 15, 8479-8499, 1987  
 A:Title: The symbiotic nitrogen fixation regulatory operon (fixRnifa) of Bradyrhizobium  
 A:Reference number: S01065; MUID:88040468; PMID:3313281  
 A:Accession: S01066  
 A:Molecule type: DNA  
 A:Residues: 1-605 <THO>  
 A:Cross-references: EMBL:X06167; NID:g39526; PIDN:CAA29531.1; PID:g39528  
 C:Genetics:  
 A:Gene: nifa  
 C:Superfamily: Response regulator (sigma54-dependent transcriptional activator), Phla ty  
 F:253-474/Domain: RNA polymerase sigma factor interaction domain homology <SFI>

Query Match 35.9%; Score 46; DB 2; Length 605;  
 Best Local Similarity 52.6%; Pred. NO. 59;  
 Matches 10; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 6 QGGGQIGDNPLLKAMKLI 24  
 ||:||||: ||:|  
 Db 249 QAHGIIGDSPALSALLEKI 267

RESULT 18  
 D75014  
 hypothetical protein PAB1257 - Pyrococcus abyssi (strain Orsay)  
 C:Species: Pyrococcus abyssi  
 C:Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 09-Jul-2004  
 C:Accession: D75014  
 R:anonymous, Genoscope  
 submitted to the EMBL Data Library, July 1999  
 A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome stru  
 A:Reference number: A75001  
 A:Accession: D75014  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-1134 <KAW>  
 A:Cross-references: UNIPROT:Q9UY58; GB:AJ248288; GB:AL096836; NID:g5458960; PIDN:CAB50505  
 A:Experimental source: strain Orsay  
 C:Genetics:  
 A:Gene: PAB1257

Query Match 35.9%; Score 46; DB 2; Length 1134;  
 Best Local Similarity 61.5%; Pred. NO. 1.2e+02;  
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 10 QIGDNPLLKAMLK 22  
 ||:||||: ||:|  
 Db 646 QVGDNPFLFSEILK 658

RESULT 19  
 T42006  
 troponin C - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
 C:Accession: T42006; T41897  
 R:Terami, H.; Williams, B.D.; Kitamura, S.; Sakube, Y.; Matsumoto, S.; Doi, S.; Obinata,  
 J. Cell Biol. 146, 193-202, 1999  
 A:Title: Genomic organization, expression, and analysis of the troponin C gene pat-10 of  
 A:Reference number: Z22021; MUID:99332308; PMID:10402470

Search completed: January 25, 2005, 12:15:59  
Job time : 15.0577 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2005, 12:07:26 ; Search time 94.2308 Seconds  
(without alignments)  
152.650 Million cell updates/sec

Title: US-09-879-248-6\_COPY\_116\_140  
Perfect score: 128  
Sequence: 1 ITPDGGGGGIGDNPLKAWLKLI 25

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : Uniprot 02.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	128	100.0	447	2	O54508 erwinia amy
2	128	100.0	447	2	O9LAW2
3	128	100.0	450	2	O6XDB8
4	128	100.0	450	2	AAQ17046
5	128	100.0	450	2	AAS45453
6	120	93.8	138	2	O79AW7
7	71	55.5	479	2	O6RK52
8	71	55.5	479	2	O6DSC8
9	71	55.5	479	2	AAS20352
10	60	46.9	237	2	O873M4
11	57	44.5	174	2	O9P923
12	56	43.8	165	2	O9P920
13	54	42.2	172	2	O9P921
14	54	42.2	220	2	O74200
15	54	42.2	837	1	STA6 MOUSE
16	53	41.4	173	2	O9P919
17	52	40.6	133	2	O49469
18	52	40.6	135	2	O53304
19	52	40.6	210	1	SODM_ASFFU
20	52	40.6	229	2	O6CFA1
21	52	40.6	584	2	O9KH13
22	52	40.6	692	2	O7NAN9
23	52	40.6	702	2	O49498
24	52	40.6	708	2	O7NAP2
25	52	40.6	710	2	O7NBR1
26	51	39.8	409	1	EFTU SYN6P
27	51	39.8	409	1	EFTU SYN7P
28	51	39.8	578	2	O7S389
29	51	39.8	667	2	O7NBR3
30	51	39.8	697	2	O7NBR5
31	50.5	39.5	504	2	O6BXD4
					O54508 erwinia amy
					O9LAW2
					O6XDB8
					AAQ17046
					AAS45453
					O79AW7
					O6RK52
					O6DSC8
					AAS20352
					O873M4
					O9P923
					O9P920
					O9P921
					O74200
					STA6 MOUSE
					O9P919
					O49469
					O53304
					SODM_ASFFU
					O6CFA1
					O9KH13
					O7NAN9
					O49498
					O7NAP2
					O7NBR1
					EFTU SYN6P
					EFTU SYN7P
					O7S389
					O7NBR3
					O7NBR5
					O6BXD4

ALIGNMENTS

O7VQG3 candidatus  
P70383 mus musculus  
O9FTW3 oryza sativ  
O7U393 synechococc  
O7VEG5 prochloroc  
O73XY5 mycobacteri  
AAS04490 mycobacte  
Q8L0W7 vibrio para  
O9RV87 deinococcoc  
Q8DIY4 synechococc  
O59360 pyrococcus  
O9FT45 arabisopsis  
O7RV88 neurospora  
Q28248 canis famil  
P50638 oryctolagus  
Q86690 canis famil  
P79197 mustela put  
P22460 homo sapien  
O7Z6M4 homo sapien  
AAS3533 homo sapi  
Q9FIE2 acetobacter  
O8IR65 drosophila  
O6RX11 human cytcm  
AAS48941 human cyt  
O8P4W4 xanthomonas  
Q8YTN7 anabaena sp  
O43133 candida alb  
Q75135 oryza sativ  
AAS7235 oryza sat  
O6RYW6 streptomyce  
AAS3904 streptomy  
Q6VMR9 myxococcus  
AAS3913 myxococcus  
Q9V9Y0 drosophila  
Q8U328 pyrococcus  
O971J4 sulfolobus  
Q9W2E1 drosophila  
Q7YU83 drosophila  
Q89PD9 bradyrhizob  
Q9XFK6 humulus lup  
Q6EV72 xenopus tro  
Q8IIM8 plasmodium  
Q748W0 geobacter s  
AAS36283 geobacter  
Q18787 caenorhabdi  
P79770 gallus gall  
Q8XW11 ralatonia s  
Q8RUT8 chlamydomon  
Q9VAM5 drosophila  
Q8T9B8 drosophila  
P43738 haemophilus  
Q89T93 bradyrhizob  
Q84T60 oryza sativ  
Q7E8Z2 oryza sativ  
BAD9637 oryza sat  
Q9V9L8 drosophila  
P27772 vibrio chol  
O67323 aquifex aeo  
Q8U5V1 agrobacteri  
Q8V9P3 streptococc  
Q8K8U6 streptococc  
Q855V7 mycobacteri  
Q866E3 alouatta be  
Q866E2 alouatta be  
Q866B9 pithecia ir  
Q866E1 alouatta be  
Q8KUX9 geobacter m  
Q8LAK9 arabisopsis  
Q9ZNT0 arabisopsis

891 2 O7VQG3  
61 2 P70383  
177 2 O9FTW3  
889 1 SYA SYNXP  
892 2 O7VEG5  
1759 2 O73XY5  
1759 2 AAS04490  
97 2 Q8L0W7  
480 2 O9RV87  
962 2 Q8DIY4  
428 2 O59360  
436 2 O9FT45  
432 2 O7RV88  
593 2 Q28248  
598 1 CIK5 RABIT  
600 2 Q866G0  
601 1 CIK5 MUSPF  
613 1 CIK5\_HUMAN  
381 2 O7Z6M4  
381 2 AAS3533  
581 2 O9FIE2  
657 2 O8IR65  
671 2 O6RX11  
671 2 AAS48941  
172 2 O8P4W4  
225 2 Q8YTN7  
238 1 TBP CANAL  
273 2 Q75135  
273 2 AAS7235  
321 2 O6RYW6  
321 2 AAS3904  
412 2 Q6VMR9  
412 2 AAS3913  
505 2 Q9V9Y0  
548 2 O8U328  
904 1 SYA SULTO  
1831 2 O9W2E1  
1872 2 Q7YU83  
262 2 Q89PD9  
316 2 Q9XFK6  
355 2 Q6EV72  
388 2 O8IIM8  
425 2 Q748W0  
425 2 AAS36283  
435 1 PRS7 CAEBEL  
443 2 P79770  
443 2 Q8XW11  
449 2 Q8RUT8  
737 2 Q9VAM5  
872 2 Q8T9B8  
874 2 Q89RG3  
901 2 Q89RG3  
1343 1 RPOB HABIN  
353 2 Q84R60  
391 2 O7E8Z2  
391 2 BAD9637  
504 2 Q9V9L8  
652 1 IRGA VIBCH  
867 1 SYA AQUAE  
55 2 Q8U5V1  
173 2 Q79YV3  
173 2 Q8K8U6  
182 2 Q855V7  
288 2 Q866E3  
290 2 Q866E2  
366 2 Q866B9  
366 2 Q866E1  
411 2 Q8KUX9  
434 2 Q8LAK9  
435 2 Q9ZNT0

32 50.5 39.5

33 50 39.1

34 50 39.1

35 50 39.1

36 50 39.1

37 50 39.1

38 50 39.1

39 49.5 38.7

40 49.5 38.7

41 49.5 38.7

42 49 38.3

43 49 38.3

44 49 38.3

45 49 38.3

46 49 38.3

47 49 38.3

48 49 38.3

49 49 38.3

50 48.5 37.9

51 48.5 37.9

52 48.5 37.9

53 48.5 37.9

54 48.5 37.9

55 48.5 37.9

56 48 37.5

57 48 37.5

58 48 37.5

59 48 37.5

60 48 37.5

61 48 37.5

62 48 37.5

63 48 37.5

64 48 37.5

65 48 37.5

66 48 37.5

67 48 37.5

68 48 37.5

69 48 37.5

70 47.5 37.1

71 47 36.7

72 47 36.7

73 47 36.7

74 47 36.7

75 47 36.7

76 47 36.7

77 47 36.7

78 47 36.7

79 47 36.7

80 47 36.7

81 47 36.7

82 47 36.7

83 47 36.7

84 46.5 36.3

85 46.5 36.3

86 46.5 36.3

87 46.5 36.3

88 46.5 36.3

89 46.5 36.3

90 46 35.9

91 46 35.9

92 46 35.9

93 46 35.9

94 46 35.9

95 46 35.9

96 46 35.9

97 46 35.9

98 46 35.9

99 46 35.9

100 46 35.9

```

RESULT 1
OS4508
ID O54508 PRELIMINARY; PRT; 447 AA.
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE HrpW protein.
GN Name=hrpW;
OS Erwinia amylovora.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Erwinia.
OX NCBI_TaxID=552;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Ea246;
RA Kim J.F., Laby R.J., Beer S.V.;
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF083620; AAF63402.1; -.
DR HSSP; Q9RHW0; IE66.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0030570; F:pectate lyase activity; IEA.
DR InterPro; IPR011050; Pectin_lyas_like.
DR InterPro; IPR004898; Pect_lyase.
DR Pfam; PF03211; Pectate_lyase; 1.
SQ SEQUENCE 447 AA; 45340 MW; 0BBAA3871EDC2F6 CRC64;

Query Match 100.0%; Score 128; DB 2; Length 447;
Best Local Similarity 100.0%; Pred. No. 5.4e-10;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ITPDGGGGGQIGDNPLKMLKLLIA 25
Db 116 ITPDGGGGGQIGDNPLKMLKLLIA 140

RESULT 3
Q6XDB8
ID O6XDB8 PRELIMINARY; PRT; 450 AA.
AC O6XDB8;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE HrpW.
GN Name=hrpW;
OS Erwinia pyrifoliae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Erwinia.
OX NCBI_TaxID=79967;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WT#3;
RA Shrestha R., Baek S.J., Bae H.N., Cho J.M., Park D.H., Hwang I.K.,
RA Hur J.H., Lim C.K.;
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=WT3;
RA Shrestha R., Baek S.J., Park D.H., Bae H.N., Cho J.M., Hwang I.,
RA Hur J.H., Lim C.K.;
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY237642; AAQ17046.1; -.
DR EMBL; AY530755; AAS45453.1; -.
DR InterPro; IPR011050; Pectin_lyas_like.
DR InterPro; IPR004898; Pect_lyase.
DR Pfam; PF03211; Pectate_lyase; 1.
SQ SEQUENCE 450 AA; 45509 MW; 085F90CD322C3A74 CRC64;

Query Match 100.0%; Score 128; DB 2; Length 450;
Best Local Similarity 100.0%; Pred. No. 5.5e-10;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ITPDGGGGGQIGDNPLKMLKLLIA 25
Db 114 ITPDGGGGGQIGDNPLKMLKLLIA 138

RESULT 4
AAQ17046
ID AAQ17046 PRELIMINARY; PRT; 450 AA.
AC AAQ17046;

```

```

RESULT 2
Q9LAW2
ID Q9LAW2 PRELIMINARY; PRT; 447 AA.
AC Q9LAW2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Harpin HrpW.

Query Match 100.0%; Score 128; DB 2; Length 447;
Best Local Similarity 100.0%; Pred. No. 5.4e-10;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ITPDGGGGGQIGDNPLKMLKLLIA 25
Db 116 ITPDGGGGGQIGDNPLKMLKLLIA 140

RESULT 3
Q6XDB8
ID O6XDB8 PRELIMINARY; PRT; 450 AA.
AC O6XDB8;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE HrpW.
GN Name=hrpW;
OS Erwinia amylovora.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Erwinia.
OX NCBI_TaxID=552;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Ea246;
RA Kim J.F., Laby R.J., Beer S.V.;
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF083620; AAF63402.1; -.
DR HSSP; Q9RHW0; IE66.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0030570; F:pectate lyase activity; IEA.
DR InterPro; IPR011050; Pectin_lyas_like.
DR InterPro; IPR004898; Pect_lyase.
DR Pfam; PF03211; Pectate_lyase; 1.
SQ SEQUENCE 447 AA; 45340 MW; 0BBAA3871EDC2F6 CRC64;

Query Match 100.0%; Score 128; DB 2; Length 447;
Best Local Similarity 100.0%; Pred. No. 5.4e-10;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ITPDGGGGGQIGDNPLKMLKLLIA 25
Db 116 ITPDGGGGGQIGDNPLKMLKLLIA 140

RESULT 4
AAQ17046
ID AAQ17046 PRELIMINARY; PRT; 450 AA.
AC AAQ17046;

```

```
DT 03-MAR-2004 (T-EMBLrel. 27, Created)
DT 03-MAR-2004 (T-EMBLrel. 27, Last sequence update)
DT 03-MAR-2004 (T-EMBLrel. 27, Last annotation update)
DE HrpW.
GN HRPW.
OS Erwinia pyrifoliae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Erwinia.
OX NCBI_TaxID=79967;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WT#3;
RA Shrestha R.; Baek S.J., Bae H.N., Cho J.M., Park D.H., Hwang I.K.,
RA Hur J.H., Lim C.K.;
RT "Molecular characterization hrp genes cluster of Erwinia pyrifoliae
RT and expression of hrpG encoding elicitor of the hypersensitive
RT response.";
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
RL EMBL; AY237642; AAQ17046.1; -.
SQ SEQUENCE 450 AA; 45509 MW; 085F90CD322C3A74 CRC64;

Query Match 100.0%; Score 128; DB 2; Length 450;
Best Local Similarity 100.0%; Pred. No. 5.5e-10;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ITPDGGGGGQIGDNPLLKAMKLIA 25
Db 114 ITPDGGGGGQIGDNPLLKAMKLIA 138

RESULT 5
AAS45453
ID AAS45453 PRELIMINARY; PRT; 450 AA.
AC AAS45453;
DT 02-MAR-2004 (T-EMBLrel. 27, Created)
DT 02-MAR-2004 (T-EMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (T-EMBLrel. 27, Last annotation update)
DE HrpW.
GN HRPW.
OS Erwinia pyrifoliae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Erwinia.
OX NCBI_TaxID=79967;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WT3;
RA Shrestha R.; Baek S.J., Park D.H., Bae H.N., Cho J.M., Hwang I.,
RA Hur J.H., Lim C.K.;
RT "Identification of hrp genes cluster and characterization of HR
RT elicitor hrpNep gene in Erwinia pyrifoliae.";
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
RL EMBL; AY530755; AAS45453.1; -.
SQ SEQUENCE 450 AA; 45509 MW; 085F90CD322C3A74 CRC64;

Query Match 100.0%; Score 128; DB 2; Length 450;
Best Local Similarity 100.0%; Pred. No. 5.5e-10;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ITPDGGGGGQIGDNPLLKAMKLIA 25
Db 114 ITPDGGGGGQIGDNPLLKAMKLIA 138

RESULT 6
Q79AW7
ID Q79AW7 PRELIMINARY; PRT; 138 AA.
AC Q79AW7;
DT 05-JUL-2004 (T-EMBLrel. 27, Created)
DT 05-JUL-2004 (T-EMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (T-EMBLrel. 27, Last annotation update)
DE HrpW (Fragment).
GN Name=hrpW;
OS Erwinia amylovora.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OX NCBI_TaxID=218491;
```

```
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Erwinia.
OX NCBI_TaxID=552;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Es321;
RX MEDLINE=98115919; PubMed=9448330;
RA Bogdanove A.J., Kim J.F., Wei Z., Kolchinsky P., Charkowski A.O.,
RA Conlin A.K., Collmer A., Beer S.V.;
RT "Homology and functional similarity of an hrp-linked pathogenicity
RT locus, dspEF, of Erwinia amylovora and the avirulence locus avrE of
RT Pseudomonas syringae pathovar tomato.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:1325-1330(1998).
DR EMBL; U97504; AAC04849.1; -.
FT NON_TER 138
SQ SEQUENCE 138 AA; 13788 MW; 4FE177177C74B3C6 CRC64;

Query Match 93.8%; Score 120; DB 2; Length 138;
Best Local Similarity 100.0%; Pred. No. 2.3e-09;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ITPDGGGGGQIGDNPLLKAMKL 23
Db 116 ITPDGGGGGQIGDNPLLKAMKL 138

RESULT 7
Q6RK52
ID Q6RK52 PRELIMINARY; PRT; 479 AA.
AC Q6RK52;
DT 05-JUL-2004 (T-EMBLrel. 27, Created)
DT 05-JUL-2004 (T-EMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (T-EMBLrel. 27, Last annotation update)
DE HrpW.
GN Name=hrpW;
OS Pectobacterium atrosepticum.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Pectobacterium.
OX NCBI_TaxID=29471;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SCRI1039;
RA Holveva M.C., Bell K.S., Hyman L.J., Avrova A.O., Whisson S.C.,
RA Birch P.R.J., Toth I.K.;
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
RL EMBL; AY496066; AAS20352.1; -.
DR InterPro; IPR011050; Pectin lyase-like.
DR InterPro; IPR004898; Pect. lyase.
DR Pfam; PF03211; Pectate lyase, 1.
SQ SEQUENCE 479 AA; 46706 MW; 3463E226CDF0406A CRC64;

Query Match 55.5%; Score 71; DB 2; Length 479;
Best Local Similarity 58.3%; Pred. No. 0.12;
Matches 14; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 1 ITPDGGGGGQIGDNPLLKAMKL 24
Db 166 INPTADGGGQLSGNDLLKALLEI 189

RESULT 8
Q6D5C8
ID Q6D5C8 PRELIMINARY; PRT; 479 AA.
AC Q6D5C8;
DT 01-OCT-2004 (T-EMBLrel. 28, Created)
DT 01-OCT-2004 (T-EMBLrel. 28, Last sequence update)
DT 01-OCT-2004 (T-EMBLrel. 28, Last annotation update)
DE Type III effector protein.
GN Name=hrpW; ORFNames=ECA2112;
OS Erwinia carotovora subsp. atroseptica SCRI1043.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Pectobacterium.
OX NCBI_TaxID=218491;
```

```

RP SEQUENCE FROM N.A.
RC STRAIN=SCRI1043;
RA Bell K.S., Sebailia M., Pritchard L., Holden M., Hyman L.J.,
RA Holeva M.C., Thomson N.R., Bentley S.D., Churcher C., Mungall K.,
RA Atkin R., Bason N., Brooks K., Chillingworth T., Clark K., Doggett J.,
RA Fraser A., Hance Z., Hauser H., Jagels K., Moule S., Norbertczak H.,
RA Ormond D., Price C., Quail M.A., Sanders M., Walker D., Whitehead S.,
RA Salmond G.P.C., Birch P.R.J., Barrell B.G., Parkhill J., Toth I.K.,
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BX950851; CAG75014.1; -.
SQ SEQUENCE 479 AA; 46706 MW; 3463E226CDF0406A CRC64;

Query Match 55.5%; Score 71; DB 2; Length 479;
Best Local Similarity 58.3%; Pred. No. 0.12;
Matches 14; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 1 ITPDGGGGGQIGDNPPLKMLKLI 24
Db 166 INPTADGGGQLSGNDLLKALLELI 189

RESULT 9
AAS20352
ID AAS20352 PRELIMINARY; PRT; 479 AA.
AC AAS20352;
DT 02-MAR-2004 (TREMBLrel. 27, Created)
DT 02-MAR-2004 (TREMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TREMBLrel. 27, Last annotation update)
DE HrpW.
GN HrpW.
OS Pectobacterium atrosepticum.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Pectobacterium.
OX NCBI_TaxID=29471;
RN [1]
RP SEQUENCE FROM N.A.
RA Holeva M.C., Bell K.S., Hyman L.J., Avrova A.O., Whisson S.C.,
RA Birch P.R.J., Toth I.K.;
RT "Rapid demonstration of a role early in disease development for the
RT type III secretion system of Erwinia carotovora subsp. atroseptica
RT SCRI1039 using a pooled transposon mutation grid.";
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY496066; AAS20352.1; -.
SQ SEQUENCE 479 AA; 46706 MW; 3463E226CDF0406A CRC64;

Query Match 55.5%; Score 71; DB 2; Length 479;
Best Local Similarity 58.3%; Pred. No. 0.12;
Matches 14; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 1 ITPDGGGGGQIGDNPPLKMLKLI 24
Db 166 INPTADGGGQLSGNDLLKALLELI 189

RESULT 10
Q873M4
ID Q873M4 PRELIMINARY; PRT; 237 AA.
AC Q873M4;
DT 01-JUN-2003 (TREMBLrel. 24, Created)
DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Manganese superoxide dismutase (Fragment).
GN Name=MnSOD;
OS Malassezia sympodialis.
OC Eukaryota; Fungi; Basidiomycota; Ustilaginomycetes;
OC Exobasidiomycetidae; Malasseziales; Malassezia.
OX NCBI_TaxID=76777;
RN [1]
RP SEQUENCE FROM N.A.
RA Andersson A., Rasool O., Schmidt M., Kodzius R., Cramer R.,
RA Scheynius A.;

```

```

RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Destroys radicals which are normally produced within the
CC cells and which are toxic to biological systems (By similarity).
CC -!- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).
CC -!- SIMILARITY: Belongs to the iron/manganese superoxide dismutase
CC family.
DR EMBL; AJ548421; CAD68071.1; -.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0004784; F:superoxide dismutase activity; IEA.
DR GO; GO:0006801; P:superoxide metabolism; IEA.
DR InterPro; IPR001189; SODismutase.
DR Pfam; PF02777; Sod_Fe_C; 1.
DR Pfam; PF00081; Sod_Fe_N; 1.
DR PRINTS; PR01703; MNSODISMUTASE.
DR PRODOM; PD000475; SODismutase; 1.
KW Oxidoreductase.
FT NON_TER 1
SQ SEQUENCE 237 AA; 26716 MW; 6C2096462E8402E9 CRC64;

Query Match 46.9%; Score 60; DB 2; Length 237;
Best Local Similarity 50.0%; Pred. No. 2.3;
Matches 11; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Qy 1 ITPDGGGGGQIGDNPPLKMLK 22
Db 118 MAPQSGGGQLNDGFLKQAIDK 139

RESULT 11
Q9P923
ID Q9P923 PRELIMINARY; PRT; 174 AA.
AC Q9P923;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Manganese-dependent superoxide dismutase (EC 1.15.1.1)
DE (Fragment).
DE Name=soda;
OS Pneumocystis carinii f. sp. muris.
OC Eukaryota; Fungi; Ascomycota; Pneumocystidomycetes; Pneumocystidaceae;
OC Pneumocystis.
OX NCBI_TaxID=42066;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20427437; PubMed=10975697;
RA Denis C.M., Mazars E., Guyot K., Odberg-Ferregut C., Viscogliosi E.,
RA Dei-Cas E., Wakefield A.E.;
RT "Genetic divergence at the SODA locus of six different formae
RT speciales of Pneumocystis carinii.";
RL Med. Mycol. 38:289-300(2000).
CC -!- FUNCTION: Destroys radicals which are normally produced within the
CC cells and which are toxic to biological systems (By similarity).
CC -!- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).
CC -!- SIMILARITY: Belongs to the iron/manganese superoxide dismutase
CC family.
DR EMBL; AF146751; AAF25723.1; -.
DR HSSP; Q92450; IKKC.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0004784; F:superoxide dismutase activity; IEA.
DR GO; GO:0006801; P:superoxide metabolism; IEA.
DR InterPro; IPR001189; SODismutase.
DR Pfam; PF02777; Sod_Fe_C; 1.
DR Pfam; PF00081; Sod_Fe_N; 1.
DR PRINTS; PR01703; MNSODISMUTASE.
DR PRODOM; PD000475; SODismutase; 1.
DR PROSITE; PS00088; SOD_MN; 1.
KW Oxidoreductase.
FT NON_TER 1
SQ SEQUENCE 174 AA; 20198 MW; 2730D8F435576124 CRC64;

```



```

OC Eukaryota; Fungi; Ascomycota; Pneumocystidomycetes; Pneumocystidaceae;
OC Pneumocystis.
OC NCBI_TaxID=42068;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=20427437; PubMed=10975697;
RA Denis C.M., Mazars E., Guyot K., Odberg-Ferragut C., Viacoglios E.,
RA Dei-Cas E., Wakefield A.E.;
RT "Genetic divergence at the SODA locus of six different formae
RT speciales of Pneumocystis carinii.";
RL Med. Mycol. 38:289-300(2000).
CC -1- FUNCTION: Destroys radicals which are normally produced within the
CC cells and which are toxic to biological systems (By similarity).
CC -1- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).
CC -1- SIMILARITY: Belongs to the iron/manganese superoxide dismutase
CC family.
DR EMBL; AF146753; AAF25725.1; -.
DR HSSP; Q92450; 1KKC.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0004784; F:superoxide dismutase activity; IEA.
DR GO; GO:0006801; P:superoxide metabolism; IEA.
DR InterPro; IPR001189; SODismutase.
DR Pfam; PF02777; Sod_Fe_C; 1.
DR Pfam; PF00081; Sod_Fe_N; 1.
DR PRINTS; PR01703; MNSODISMUTASE.
DR ProDom; PD000475; SODismutase. 1.
DR PROSITE; PS00088; SOD_MN; 1.
DR KX Oxidoreductase.
DR FT NON_TER 1 1
DR FT NON_TER 172 172
DR SQ SEQUENCE 172 AA; 19894 MW; CDEC754E7F7F8F2A CRC64;

Query Match 42.2%; Score 54; DB 2; Length 172;
Best Local Similarity 45.5%; Pred. No. 12;
Matches 10; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 1 ITPDGGGGGQIGDNPLLKAWLK 22
: : : : : : : : : : : : : :
DB 69 LLPPKEGGGQVLDGPLVDIAIKK 90

RESULT 14
O74200 PRELIMINARY; PRT; 220 AA.
ID 074200 AC
AC O74200;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Manganese superoxide dismutase precursor.
DE Name=msod;
DE OS Pneumocystis carinii.
DE OC Eukaryota; Fungi; Ascomycota; Pneumocystidomycetes; Pneumocystidaceae;
OC Pneumocystis.
OC NCBI_TaxID=4754;
RN [1]
RN SEQUENCE FROM N.A.
RA Denis C.M., Guyot K., Dei-Cas E., Camus D., Odberg-Ferragut C.;
RA Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
RL -1- FUNCTION: Destroys radicals which are normally produced within the
RL cells and which are toxic to biological systems (By similarity).
RL -1- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).
CC -1- SIMILARITY: Belongs to the iron/manganese superoxide dismutase
CC family.
DR EMBL; AF036321; AAC24764.1; -.
DR HSSP; Q92450; 1KKC.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0004784; F:superoxide dismutase activity; IEA.
DR GO; GO:0006801; P:superoxide metabolism; IEA.
DR InterPro; IPR000834; Peptidase_M14.
DR InterPro; IPR001189; SODismutase.
DR Pfam; PF02777; Sod_Fe_C; 1.

```

```
DR Pfam; PF00081; Sod_Fe_N; 1.
DR PRINTS; PR01703; MNSODISMUTASE.
DR ProDom; PD000475; SODISMUTASE; 1.
DR PROSITE; PS00133; CARBOXYPEPT_ZN_2; UNKNOWN_1.
DR PROSITE; PS00088; SOD_MN; 1.
KW Oxidoreductase; Signal.
FT SIGNAL 1 27 Potential.
FT CHAIN 48 169 manganese superoxide dismutase.
FT CHAIN 170 183 manganese superoxide dismutase.
FT CHAIN 184 207 manganese superoxide dismutase.
FT CHAIN 208 220 manganese superoxide dismutase.
FT CHAIN 28 35 manganese superoxide dismutase.
FT CHAIN 36 47 manganese superoxide dismutase.
SQ SEQUENCE 220 AA; 25869 MM; 73B1F11C98929E18 CRC64;

Query Match 42.2%; Score 54; DB 2; Length 220;
Best Local Similarity 45.5%; Pred. No. 16;
Matches 10; Conservative 7; Mismatches 0; Gaps 0;

QY 1 ITPDGGGGGQIGDNPFLKMLK 22
: | : ||| : | : ||| :
Db 106 LLPNKQGGGEIINGPLVEAIKK 127

RESULT 15
STA6 MOUSE STANDARD; PRT; 837 AA.
ID STA6 MOUSE STANDARD; PRT; 837 AA.
AC P52633;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Signal transducer and transcription activator 6.
GN Name=Stat6;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95280934; PubMed=7760829;
RA Quelle F.W., Shmida K., Thierfelder W., Fischer C.L., Kim A.,
RA Ruben S.M., Cleveland J.L., Pierce J.H., Keegan A.D., Nelms K.,
RA Paul W.E., Ihle J.N.;
RT "Cloning of murine Stat6 and human Stat6, Stat proteins that are
RT tyrosine phosphorylated in responses to IL-4 and IL-3 but are not
RT required for mitogenesis";
RL Mol. Cell. Biol. 15:3336-3343(1995).
CC -!- FUNCTION: Carries out a dual function: signal transduction and
CC activation of transcription. Involved in interleukin-4 signalling.
CC -!- SUBUNIT: Forms a homodimer or a heterodimer with a related family
CC member (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic; translocated into the nucleus
CC in response to phosphorylation.
CC -!- PTM: Tyrosine phosphorylated following stimulation by IL-4 and IL-
CC 3.
CC -!- SIMILARITY: Belongs to the transcription factor STAT family.
CC -!- SIMILARITY: Contains 1 SH2 domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L47650; AAN79006.1; -.
DR PIR; I57557; I57557.
DR HSSP; P42227; IBGI.
DR MGD; MGI:103034; Stat6.
DR GO; GO:0005737; C:cytoplasm; IDA.
DR GO; GO:0005634; C:nucleus; IDA.
DR GO; GO:0042127; P:regulation of cell proliferation; IDA.
```

```
DR InterPro; IPR008967; P53_like_DNA_bnd.
DR InterPro; IPR000980; SH2.
DR InterPro; IPR001217; STAT.
DR Pfam; PF00017; SH2; 1.
DR Pfam; PF01017; STAT_alpha; 1.
DR Pfam; PF02864; STAT_bind; 1.
DR Pfam; PF02865; STAT_int; 1.
DR SMART; SM00252; SH2; 1.
DR PROSITE; PS00001; SH2; 1.
KW Activator; DNA-binding; Nuclear protein; Phosphorylation; SH2 domain;
KW Transcription regulation.
FT DOMAIN 517 632 SH2.
FT MOD RES 641 641 Phosphotyrosine (by JAK) (By similarity).
SQ SEQUENCE 837 AA; 93725 MM; FD7D7C48743EFC7 CRC64;

Query Match 42.2%; Score 54; DB 1; Length 837;
Best Local Similarity 64.3%; Pred. No. 64;
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 5 GGGGGQIGDNPFLK 18
: ||| : |||
Db 802 GGGGSLGSQLK 815
```

```
RESULT 16
Q9P919 PRELIMINARY; PRT; 173 AA.
ID Q9P919 PRELIMINARY; PRT; 173 AA.
AC Q9P919;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Manganese-dependent superoxide dismutase (EC 1.15.1.1)
DE (Fragment).
DE Name=soda;
OS Pneumocystis carinii f. sp. suis.
OC Eukaryota; Fungi; Ascomycota; Pneumocystidomycetes; Pneumocystidaceae;
OC Pneumocystis.
OX NCBI_TaxID=112251;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20427437; PubMed=10975697;
RA Denis C.M., Mazars E., Guyot K., Odberg-Ferragut C., Viscogliosi E.,
RA Del-Cas E., Wakefield A.E.;
RT "Genetic divergence at the SODA locus of six different formae
RT specialises of Pneumocystis carinii.";
RL Med. Mycol. 38:289-300(2000).
CC -!- FUNCTION: Destroys radicals which are normally produced within the
CC cells and which are toxic to biological systems (By similarity).
CC -!- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).
CC -!- SIMILARITY: Belongs to the iron/manganese superoxide dismutase
CC family.
DR EMBL; AF146755; AAF25727.1; -.
DR HSSP; Q92450; 1KKC.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0004784; F:superoxide dismutase activity; IEA.
DR GO; GO:0006801; P:superoxide metabolism; IEA.
DR InterPro; IPR001189; SODismutase.
DR Pfam; PF02777; Sod_Fe_C; 1.
DR Pfam; PF00081; Sod_Fe_N; 1.
DR PRINTS; PR01703; MNSODISMUTASE.
DR ProDom; PD000475; SODISMUTASE; 1.
DR PROSITE; PS00088; SOD_MN; 1.
KW Oxidoreductase.
FT NON TER 1
FT NON TER 173 173
SQ SEQUENCE 173 AA; 20099 MM; 69110EFD0005665B CRC64;

Query Match 41.4%; Score 53; DB 2; Length 173;
Best Local Similarity 40.9%; Pred. No. 17;
Matches 9; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

QY 1 ITPDGGGGQIGDNPFLKMLK 22
```

Db 70 LILPQKGGGKVDGFLVEAIKK 91

RESULT 17

Q49469 PRELIMINARY; PRT; 133 AA.

AC Q49469; 133 AA.

DT 01-NOV-1996 (TrEMBLrel. 01, Created)

DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DE Haemagglutinin homologue precursor (Fragment).

GN Name=pmGAI.3;

OS Mycoplasma gallisepticum.

OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.

ON NCBI\_TaxID=2096;

OX [1]

RP SEQUENCE FROM N.A.

RA Markham P.F.;

RL Submitted (NOV-1992) to the EMBL/GenBank/DBJ databases.

DR EMBL; M83178; AAA02997.1; --.

KW SIGNAL

FT CHAIN 1 25 Potential.

FT NON\_TER 26 >133 haemagglutinin homologue.

FT SQ SEQUENCE 133 AA; 13628 MW; 612769FDB8265A496 CRC64;

Query Match 40.6%; Score 52; DB 2; Length 133;

Best Local Similarity 64.3%; Pred. No. 18;

Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 TPDGGGGGGQIGDNP 15

DB 57 TNPNGGGGMMGDNP 70

RESULT 18

Q53304 PRELIMINARY; PRT; 135 AA.

AC Q53304; 135 AA.

DT 01-NOV-1996 (TrEMBLrel. 01, Created)

DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DE Haemagglutinin homologue (Fragment).

GN Name=pmGAI.3;

OS Mycoplasma gallisepticum.

OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.

ON NCBI\_TaxID=2096;

OX [1]

RP SEQUENCE FROM N.A.

RA Markham P.F.; Glew M.D., Whithear K.G., Walker I.D.;

RL Molecular cloning of a member of the gene family that encodes pmGA, a hemagglutinin of Mycoplasma gallisepticum.

RT Infect. Immun. 61:903-909 (1993).

DR EMBL; S55216; AAB25398.2; --.

DR PIR; B49218; B49218.

FT NON\_TER 135

FT SQ SEQUENCE 135 AA; 13826 MW; 9F77012769FDB8265 CRC64;

Query Match 40.6%; Score 52; DB 2; Length 135;

Best Local Similarity 64.3%; Pred. No. 18;

Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 TPDGGGGGGQIGDNP 15

DB 57 TNPNGGGGMMGDNP 70

RESULT 19

SODM ASPFU STANDARD; PRT; 210 AA.

AC Q92450;

Db 70 LILPQKGGGKVDGFLVEAIKK 91

RESULT 17

Q49469 PRELIMINARY; PRT; 133 AA.

AC Q49469; 133 AA.

DT 01-NOV-1996 (TrEMBLrel. 01, Created)

DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DE Haemagglutinin homologue precursor (Fragment).

GN Name=pmGAI.3;

OS Mycoplasma gallisepticum.

OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.

ON NCBI\_TaxID=2096;

RN [1]

RP SEQUENCE FROM N.A.

RA Markham P.F.;

RL Submitted (NOV-1992) to the EMBL/GenBank/DBJ databases.

DR EMBL; M83178; AAA02997.1; --.

KW SIGNAL.

FT SIGNAL 1 25 Potential.

FT CHAIN 26 >133 haemagglutinin homologue.

FT NON\_TER 133

FT SEQUENCE 133 AA; 13628 MW; 612769FD8265A496 CRC64;

SQ

Query Match 40.6%; Score 52; DB 2; Length 133;

Best Local Similarity 64.3%; Pred. No. 18;

Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 TPDGGGGGQIGDNP 15

DB 57 TNPNGGGMMGDNP 70

RESULT 18

Q53304 PRELIMINARY; PRT; 135 AA.

AC Q53304; 135 AA.

DT 01-NOV-1996 (TrEMBLrel. 01, Created)

DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DE Haemagglutinin homologue (Fragment).

GN Name=pmGAI.3;

OS Mycoplasma gallisepticum.

OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.

ON NCBI\_TaxID=2096;

RN [1]

RP SEQUENCE FROM N.A.

RA Markham P.F.; Glew M.D., Whithear K.G., Walker I.D.;

RL MEDLINE=93162830; PubMed=8432610;

RT Molecular cloning of a member of the gene family that encodes pmGA, a hemagglutinin of Mycoplasma gallisepticum.;

RT Infect. Immun. 61:903-909 (1993).

RL EMBL; S55216; AAB25398.2; --.

DR PIR; B49218; B49218.

FT NON\_TER 135

FT SEQUENCE 135 AA; 13826 MW; 9F77012769FD8265 CRC64;

SQ

Query Match 40.6%; Score 52; DB 2; Length 135;

Best Local Similarity 64.3%; Pred. No. 18;

Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 TPDGGGGGQIGDNP 15

DB 57 TNPNGGGMMGDNP 70

RESULT 19

SODM ASPFU STANDARD; PRT; 210 AA.

AC Q92450;



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2005, 12:06:11 ; Search time 60.75 Seconds  
(without alignments)  
159.435 Million cell updates/sec

residue 116-140

Title: us-09-879-248-6\_copy\_31\_57  
Perfect score: 128  
Sequence: 1 NSALGQPDRTQIEQMAQLLAKS 27

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0  
Maximum DB seq length: 200000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 100 summaries

Database : A\_Geneseq\_23Sep04:\*

- 1: Geneseq1980s:\*
- 2: Geneseq1990s:\*
- 3: Geneseq2000s:\*
- 4: Geneseq2001s:\*
- 5: Geneseq2002s:\*
- 6: Geneseq2003as:\*
- 7: Geneseq2003bs:\*
- 8: Geneseq2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	128	100.0	447	2	AAW96260 Hypersens
2	128	100.0	447	3	AAW71094 Erwinia a
3	128	100.0	447	3	AAW84855 A hyperse
4	128	100.0	447	5	AAO22548 Hypersens
5	128	100.0	447	5	AAE18296 Erwinia a
6	128	100.0	447	5	AAE16448 E. amylov
7	128	100.0	447	5	ABO92226 Erwinia a
8	52	40.6	313	7	ABO73754 Pseudomon
9	52	40.6	457	7	ABO73936 Pseudomon
10	52	40.6	710	8	ADN73161 Thale cre
11	51	39.8	660	7	ADF29258 Bacterial
12	51	39.8	660	7	ADG32467 Pasteurel
13	51	39.8	1312	2	AAW22775 Human RAD
14	51	39.8	1312	2	AAW71295 Human hom
15	51	39.8	1318	5	ABW77985 Amino aci
16	51	39.8	1318	7	ADJ68860 Human hea
17	51	39.8	1318	8	ADJ66509 RAD50 hom
18	50.5	39.5	286	5	ABG70071 Human pre
19	50.5	39.5	352	2	AAW22307 Lipase mo
20	50.5	39.5	596	5	ABG61825 Prostate
21	50.5	39.5	596	7	ABW75398 Prostate
22	50	39.1	430	5	ABW47515 Listeria
23	49	38.3	264	7	ABO71596 Pseudomon
24	49	38.3	697	6	ABR40712 Oryza sat
25	49	38.3	3249	6	ABU39648 Protein e

ABO70264	Pseudomon	26	48.5	37.9	237	7	ABO70264
ADC03511	Pseudomon	27	48.5	37.9	777	7	ADC03511
ABO76849	Pseudomon	28	48.5	37.9	891	7	ABO76849
ABG60921	Homo sapi	29	48	37.5	223	4	ABG60921
ADH13660	Human ENZ	30	48	37.5	320	8	ADH13660
ABP59216	Human dru	31	48	37.5	422	6	ABP59216
ABG69596	Human NOV	32	48	37.5	492	5	ABG69596
ADJ83019	Human NOV	33	48	37.5	492	7	ADJ83019
ADJ83077	Human cyt	34	48	37.5	495	6	ADJ83077
ABU21525	Protein e	35	48	37.5	630	6	ABU21525
AAC020572	Protein E	36	48	37.5	1419	5	AAC020572
ABG99408	Amino aci	37	48	37.5	1419	6	ABG99408
ADN48494	Human smg	38	48	37.5	1419	8	ADN48494
ADJ70004	Human hea	39	48	37.5	1449	7	ADJ70004
ABU41616	Protein e	40	48	37.5	2246	6	ABU41616
ABG63216	Klebsiell	41	47	36.7	112	7	ABG63216
ADN47050	Thermococ	42	47	36.7	183	8	ADN47050
ADN97661	S ambofac	43	47	36.7	301	8	ADN97661
AP80242	Sequence	44	47	36.7	320	1	AP80242
AP80714	Sequence	45	47	36.7	320	1	AP80714
ADN97691	S ambofac	46	47	36.7	387	8	ADN97691
ABO74869	Pseudomon	47	47	36.7	505	7	ABO74869
ADQ11466	Human Y2	48	46	35.9	31	8	ADQ11466
ADQ11465	Human Y2	49	46	35.9	32	8	ADQ11465
ADQ11464	Human Y2	50	46	35.9	33	8	ADQ11464
ADQ11463	Human Y2	51	46	35.9	34	8	ADQ11463
ABR98024	Human pan	52	46	35.9	36	6	ABR98024
ABR98049	Bovine pa	53	46	35.9	36	6	ABR98049
ADBE2966	Human pan	54	46	35.9	36	7	ADBE2966
ADBE2991	Bovine pa	55	46	35.9	36	7	ADBE2991
ADEI2880	Human pan	56	46	35.9	36	7	ADEI2880
ADEI2885	Cpw pancr	57	46	35.9	36	7	ADEI2885
ADQ11461	Human Y2	58	46	35.9	36	8	ADQ11461
ABP01232	Human ORF	59	46	35.9	72	5	ABP01232
ADN47991	Thermococ	60	46	35.9	158	8	ADN47991
ABW78891	C. glutam	61	46	35.9	449	4	ABW78891
AGS2396	C. glutam	62	46	35.9	449	4	AGS2396
ABP81013	N. gonorr	63	45.5	35.5	249	6	ABP81013
ABU37237	Protein e	64	45.5	35.5	249	6	ABU37237
ABU37734	Protein e	65	45.5	35.5	249	6	ABU37734
ADP08371	Neisseria	66	45.5	35.5	249	8	ADP08371
ABU23294	Protein e	67	45.5	35.5	422	6	ABU23294
ADQ11468	Human Y2	68	45	35.2	29	8	ADQ11468
ADQ11467	Human Y2	69	45	35.2	30	8	ADQ11467
ABR98048	Cat pancr	70	45	35.2	36	6	ABR98048
ADBE2990	Cat pancr	71	45	35.2	36	7	ADBE2990
ADBE2884	Cat Y2 re	72	45	35.2	36	7	ADBE2884
ADQ11499	Cat Y2 re	73	45	35.2	36	8	ADQ11499
AY05276	Pancrati	74	45	35.2	95	2	AY05276
ABU02140	S. pneumo	75	45	35.2	203	6	ABU02140
AGS91339	C. glutam	76	45	35.2	252	4	AGS91339
ADC59410	P. pneumo	77	45	35.2	307	7	ADC59410
ADG64963	Protein #	78	45	35.2	307	7	ADG64963
AU34354	Staphyloc	79	45	35.2	489	4	AU34354
ABO71672	Pseudomon	80	45	35.2	631	7	ABO71672
AAW60664	Human ALA	81	45	35.2	756	2	AAW60664
ABU16227	Protein e	82	45	35.2	978	6	ABU16227
AU36618	Staphyloc	83	45	35.2	981	6	AU36618
ABW71093	Staphyloc	84	45	35.2	981	6	ABW71093
AAW24559	Presenili	85	45	35.2	1040	2	AAW24559
AY23900	Human res	86	45	35.2	1084	2	AY23900
ABO7973	A human c	87	45	35.2	1225	3	ABO7973
ADF09504	Human cat	88	45	35.2	1225	7	ADF09504
ABG04996	Novel hum	89	45	35.2	1233	4	ABG04996
ABO7974	A murine	90	45	35.2	1247	3	ABO7974
ABU36321	Protein e	91	45	35.2	1818	6	ABU36321
ABG84277	Adenoviru	92	45	35.2	11300	5	ABG84277
ABG23536	Arabicdops	93	44.5	34.8	250	3	ABG23536
AG27969	Arabicdops	94	44.5	34.8	397	3	AG27969
AG46302	Arabicdops	95	44.5	34.8	397	3	AG46302
ADC00859	Enterohae	96	44.5	34.8	454	7	ADC00859
ADC00483	Enterohae	97	44.5	34.8	833	7	ADC00483
ADC00331	Enterohae	98	44.5	34.8	859	7	ADC00331

99 44.5 34.8 859 7 ADC00051 Enterohae  
100 44.5 34.8 1153 6 ABU33426 Protein e

# ALIGNMENTS

RESULT 1  
AAW96260  
ID AAW96260 standard; protein; 447 AA.  
AC AAW96260;  
XX  
DT 14-JUN-1999 (first entry)  
XX  
DE Hypersensitive response eliciting protein (HrpW).  
XX  
KW Hypersensitive response elicitor protein; hypersensitive response; hrpW;  
KW pathogen; infection; crop protection; disease resistance;  
KW pest resistance; transgenic plant; colouration; maturation.  
XX  
OS Erwinia amylovora.  
XX  
PN WO9907208-A1.  
XX  
PD 18-FEB-1999.  
XX  
PF 27-JUL-1998; 98WO-US015547.  
XX  
PR 06-AUG-1997; 97US-0055108P.  
XX  
PA (CORR ) CORNELL RES FOUND INC.  
XX  
PI Kim JF, Beer SV;  
XX  
DR WPI; 1999-167126/14.  
XX  
DR N-PSDB; AAX09007.  
XX

PT New Erwinia amylovora hypersensitive response eliciting gene and protein  
PT - useful for providing transgenic plants and seeds with enhanced growth,  
PT and insect and disease resistance.  
XX  
PS Claim 1; Page 50-51; 54pp; English.  
XX  
CC The hypersensitive response eliciting protein (hrp) or polypeptide is  
CC produced as part of an active defense by plants against incompatible  
CC pathogen infections. The hypersensitive response is a rapid localised  
CC necrosis. The hrp protein and gene when used in nucleotide constructs are  
CC useful for providing disease resistance to plants, insect control to  
CC plants, and enhancing plant growth (enhancing fruit size and earlier  
CC colouration and maturation), by direct application of the protein to  
CC plants, or by producing transgenic plants or seeds using the hrp gene  
XX  
SQ Sequence 447 AA;  
XX

Query Match 100.0%; Score 128; DB 2; Length 447;  
Best Local Similarity 100.0%; Pred. No. 6.4e-11;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 NSALGQOPIDRQTIEQMAQLLAELLS 27  
DB 31 NSALGQOPIDRQTIEQMAQLLAELLS 57  
RESULT 2  
AAV71094  
ID AAV71094 standard; protein; 447 AA.  
AC AAV71094;  
XX  
DT 08-SEP-2000 (first entry)  
XX  
DE Erwinia amylovora hypersensitive response elicitor #2.

Query Match 100.0%; Score 128; DB 2; Length 447;  
Best Local Similarity 100.0%; Pred. No. 6.4e-11;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 NSALGQOPIDRQTIEQMAQLLAELLS 27  
DB 31 NSALGQOPIDRQTIEQMAQLLAELLS 57

RESULT 3  
AAV71094  
ID AAV71094 standard; protein; 447 AA.  
AC AAV71094;  
XX  
DT 08-AUG-2000 (first entry)  
XX  
DE A hypersensitive response elicitor protein.  
XX  
KW Hypersensitive response; insect control; disease resistance;  
KW hypersensitive response elicitor; plant growth; vegetable; crop;  
KW ornamental plant.  
XX  
OS Erwinia amylovora.  
XX  
PN WO200020452-A2.  
XX  
PD 13-APR-2000.  
XX  
PF 05-OCT-1999; 99WO-US023181.  
XX  
PR 05-OCT-1998; 98US-0103050P.  
XX  
PA (EDEN-) EDEN BIOSCIENCE CORP.  
XX  
PI Wei Z, Fan H, Niggemeyer JL;  
XX  
DR WPI; 2000-303745/26.  
XX  
DR N-PSDB; AAA14939.

XX Hypersensitive response elicitor; environmental stress resistance; plant.  
KW Erwinia amylovora.  
OS  
XX WO200028055-A2.  
XX  
PD 18-MAY-2000.  
XX  
PF 04-NOV-1999; 99WO-US026039.  
XX  
PR 05-NOV-1998; 98US-0107243P.  
XX  
PA (EDEN-) EDEN BIOSCIENCE CORP.  
XX  
PI Wei Z, Schading RL;  
XX  
DR WPI; 2000-376566/32.  
XX  
DR N-PSDB; AAD00669.  
XX  
PT Application of a hypersensitive response elicitor protein to plants to  
PT impart stress resistance.  
XX  
PS Disclosure; Page 10-12; 84pp; English.  
XX  
CC The patent discloses a method to impart stress resistance to plants by  
CC applying a hypersensitive response elicitor in a non-infectious form to a  
CC plant or seed. The present sequence is a hypersensitive response elicitor  
CC protein from Erwinia amylovora. The protein is heat stable, protease  
CC sensitive and suppressed by inhibitors of plant metabolism. It is used to  
CC impart stress resistance to plants  
XX  
SQ Sequence 447 AA;

Query Match 100.0%; Score 128; DB 3; Length 447;  
Best Local Similarity 100.0%; Pred. No. 6.4e-11;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 NSALGQOPIDRQTIEQMAQLLAELLS 27  
DB 31 NSALGQOPIDRQTIEQMAQLLAELLS 57

RESULT 3  
AAV71094  
ID AAV71094 standard; protein; 447 AA.  
XX  
AC AAV71094;  
XX  
DT 08-AUG-2000 (first entry)  
XX  
DE A hypersensitive response elicitor protein.  
XX  
KW Hypersensitive response; insect control; disease resistance;  
KW hypersensitive response elicitor; plant growth; vegetable; crop;  
KW ornamental plant.  
XX  
OS Erwinia amylovora.  
XX  
PN WO200020452-A2.  
XX  
PD 13-APR-2000.  
XX  
PF 05-OCT-1999; 99WO-US023181.  
XX  
PR 05-OCT-1998; 98US-0103050P.  
XX  
PA (EDEN-) EDEN BIOSCIENCE CORP.  
XX  
PI Wei Z, Fan H, Niggemeyer JL;  
XX  
DR WPI; 2000-303745/26.  
XX  
DR N-PSDB; AAA14939.

XX Hypersensitive response elicitor polypeptides useful for imparting  
PT enhanced growth, disease resistance and insect resistance to plants,  
PT especially vegetables and ornamental flowers.  
XX  
XX PS Disclosure; Page 12-13; 100pp; English.  
XX  
XX The present sequence represents a hypersensitive response elicitor  
CC polypeptide. The specification describes hypersensitive response elicitor  
CC polypeptide fragments, which do not elicit a hypersensitive response.  
CC Instead, the proteins impart disease resistance to plants, enhance plant  
CC growth, and/or control insects. The polypeptide fragments may be used to  
CC these properties to plants. The plants which may be treated in this way  
CC include vegetables, crops and ornamental plants such as alfalfa, rice,  
CC wheat, barley, rye, cotton, sunflower, peanut, corn, potato, sweet  
CC potato, bean, pea, chickory, lettuce, endive, cabbage, brussel sprout,  
CC beet, parsnip, turnip, cauliflower, broccoli, radish, spinach, onion,  
CC garlic, eggplant, pepper, celery, carrot, squash, pumpkin, zucchini,  
CC cucumber, apple, pear, melon, citrus, strawberry, grape, raspberry,  
CC pineapple, soybean, tobacco, tomato, sorghum or sugarcane, Arabidopsis  
CC thaliana, Saintpaulia, petunia, pelargonium, poinsettia, chrysanthemum,  
CC carnation or zinnia  
XX  
XX Sequence 447 AA;  
Query Match 100.0%; Score 128; DB 3; Length 447;  
Best Local Similarity 100.0%; Pred. No. 6.4e-11;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 NSALGQGPIDRQTIEQMAQLLAELLKS 27  
DB 31 NSALGQGPIDRQTIEQMAQLLAELLKS 57  
RESULT 4  
AAO22548  
ID AAO22548 standard; protein; 447 AA.  
XX  
XX AAO22548;  
XX  
XX 28-OCT-2002 (first entry)  
XX  
XX Hypersensitive response elicitor protein #2 from Erwinia amylovora.  
XX  
XX Desiccation inhibitor; longevity enhancer; desiccation inhibitor;  
KW cutting; ornamental plant; hypersensitive response elicitor protein;  
KW transgenic ornamental plant; early flowering; longevity enhancer; plant;  
KW flower bloom; flower.  
XX  
XX Erwinia amylovora.  
OS  
XX WO200237960-A2.  
XX  
XX 16-MAY-2002.  
XX  
XX 06-NOV-2001; 2001WO-US043715.  
XX  
XX 13-NOV-2000; 2000US-0248169P.  
XX  
XX (EDEN-) EDEN BIOSCIENCE CORP.  
XX  
XX Wei Z, Leon E, Oviedo A;  
PI WPI; 2002-575194/61.  
DR N-PSDB; AAL41133.  
XX  
XX Inhibiting desiccation of cuttings from ornamental plants, by treating  
PT ornamental plants with hypersensitive response elicitor protein, or  
PT expressing heterologous hypersensitive response elicitor protein in  
PT plants.  
XX  
XX Disclosure; Page 12-13; 69pp; English.  
XX

CC The invention relates to a method for inhibiting desiccation of cuttings  
CC from ornamental plants. The method involves treating the cuttings with a  
CC hypersensitive response elicitor protein or polypeptide, or providing a  
CC transgenic ornamental plant or plant seed transformed with a DNA molecule  
CC encoding the hypersensitive response elicitor polypeptide, and growing  
CC the ornamental plant or transgenic ornamental plant produced from the  
CC transgenic ornamental plant seed. The hypersensitive response elicitor  
CC protein or polypeptide is useful for inhibiting desiccation of cuttings  
CC from ornamental plants, for harvesting cuttings from ornamental plants,  
CC for promoting early flowering of ornamental plants, and enhancing the  
CC longevity of flower blooms on ornamental plant cuttings. This sequence  
CC represents a hypersensitive response elicitor protein of the invention  
XX  
XX Sequence 447 AA;  
Query Match 100.0%; Score 128; DB 5; Length 447;  
Best Local Similarity 100.0%; Pred. No. 6.4e-11;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 NSALGQGPIDRQTIEQMAQLLAELLKS 27  
DB 31 NSALGQGPIDRQTIEQMAQLLAELLKS 57  
RESULT 5  
AAE18296  
ID AAE18296 standard; protein; 447 AA.  
XX  
XX AAE18296;  
XX  
XX 07-MAY-2002 (first entry)  
XX  
XX Erwinia amylovora hypersensitive response elicitor (HRE) #2.  
XX  
XX Hypersensitive response elicitor; HRE; transgenic plant; plant growth;  
KW stress tolerance; disease tolerance; modified flower colour;  
KW insect resistance; herbicide resistance; male sterility.  
XX  
XX Erwinia amylovora.  
OS  
XX WO200195724-A2.  
XX  
XX 20-DEC-2001.  
XX  
XX 13-JUN-2001; 2001WO-US018955.  
XX  
XX 15-JUN-2000; 2000US-0211585P.  
XX  
XX (EDEN-) EDEN BIOSCIENCE CORP.  
XX  
XX Wei Z, Derocher J;  
PI WPI; 2002-130707/17.  
DR N-PSDB; AAD29125.  
XX  
XX Improving effectiveness of transgenic plants by topical application of a  
PT hypersensitive response elicitor protein to the transgenic plant or by  
PT incorporating into the plant a transgene encoding the protein.  
XX  
XX Disclosure; Page 13-14; 86pp; English.  
XX  
XX The invention relates to methods of improving the effectiveness of  
CC transgenic plants which involves either topical application of a  
CC hypersensitive response elicitor (HRE) protein to the transgenic plant or  
CC incorporating into the transgenic plant a transgene encoding HRE. HRE  
CC sequence is used for improving the effectiveness of transgenic plants by  
CC maximising the benefit of transgenic traits associated with a deleterious  
CC effect on growth, stress tolerance, disease or insect resistance,  
CC enhanced growth, herbicide resistance, male sterility, modified flower  
CC colour and biochemically modified plant product in the transgenic plants  
CC or overcoming the deleterious effects. The present sequence is Erwinia  
CC amylovora HRE protein  
XX

SQ Sequence 447 AA;

Query Match 100.0%; Score 128; DB 5; Length 447;  
 Best Local Similarity 100.0%; Pred. No. 6,4e-11;  
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSALGQOPIDRQTIEQMAQLLAELLS 27  
 |||||  
 Db 31 NSALGQOPIDRQTIEQMAQLLAELLS 57

RESULT 6  
 AAE16448  
 ID AAE16448 standard; protein; 447 AA.  
 XX  
 AC AAE16448;  
 XX  
 DT 09-APR-2002 (first entry)  
 XX  
 DE E. amylovora hypersensitive response elicitor protein #2.  
 XX  
 KW Hypersensitive response elicitor protein; plant growth; fruit coloration;  
 KW disease resistance; stress resistance; phytoalexin; insect infection;  
 KW plant maturation.  
 XX  
 OS Erwinia amylovora.

Key Location/Qualifiers  
 FT Domain 5..64  
 FT /label= Hypersensitive\_response\_eliciting\_domain\_1  
 FT Region 5..45  
 FT /label= Acidic\_unit  
 FT Region 45..64  
 FT /label= Alpha\_helix  
 FT Domain 103..146  
 FT /label= Hypersensitive\_response\_eliciting\_domain\_2  
 FT Region 103..131  
 FT /label= Acidic\_unit  
 FT Region 131..146  
 FT /label= Alpha\_helix  
 XX  
 PN WO200198501-A2.  
 XX  
 PD 27-DEC-2001.  
 XX  
 PF 12-JUN-2001; 2001WO-US018820.  
 XX  
 PR 16-JUN-2000; 2000US-0212211P.  
 XX  
 PA (EDEN-) EDEN BIOSCIENCE CORP.  
 XX  
 PI Fan H, Wei Z;  
 XX  
 WI 2002-122282/16.  
 DR N-PSDB; AAD27016.  
 XX  
 PT New hypersensitive response elicitor proteins comprising spaced apart  
 PT domains having an acidic portion linked to an alpha-helix, useful for  
 PT imparting disease or stress resistance, controlling insects or enhancing  
 PT plant growth.  
 XX  
 PS Disclosure; Page 13-14; 99pp; English.  
 XX  
 CC The patent discloses hypersensitive response elicitor proteins and  
 CC nucleotides encoding such proteins. Hypersensitive response elicitor  
 CC proteins comprise an isolated pair or more of spaced apart domains, each  
 CC comprising an acidic portion linked to an alpha-helix and capable of  
 CC eliciting a hypersensitive response in plants. Sequences of the invention  
 CC are used to impart disease resistance to plants, to enhance plant growth,  
 CC to control insects and/or to impart stress resistance to plants which  
 CC includes resistance to environmental stresses such as climate, air  
 CC pollution, chemical and nutritional stress. The method of imparting  
 CC disease resistance has the potential for treating previously untreatable

diseases, treating diseases systemically and avoiding the use of  
 infectious agents or environmentally harmful materials. Hyper- sensitive  
 response elicitor sequences are used to enhance plant growth which  
 encompasses greater yield, increased in quantity of seeds produced,  
 percentage of seeds germinated, plant size and biomass, bigger fruits,  
 earlier fruit coloration and plant maturation. They are also used for  
 insect control which encompasses preventing direct insect damage to plant  
 by feeding injury, interfering with insect larval feeding on the plants,  
 preventing insects from colonizing host plants and releasing phytotoxins.  
 Sequences of the invention also prevent subsequent disease damage to  
 plants resulting from insect infection. The present sequence is Erwinia  
 amylovora hypersensitive response elicitor protein

SQ Sequence 447 AA;

Query Match 100.0%; Score 128; DB 5; Length 447;  
 Best Local Similarity 100.0%; Pred. No. 6,4e-11;  
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSALGQOPIDRQTIEQMAQLLAELLS 27  
 |||||  
 Db 31 NSALGQOPIDRQTIEQMAQLLAELLS 57

RESULT 7  
 ABB09226  
 ID ABB09226 standard; protein; 447 AA.  
 XX  
 AC ABB09226;  
 XX  
 DT 08-JUL-2002 (first entry)  
 XX  
 DE Erwinia amylovora hypersensitive response elicitor SEQ ID NO:5.  
 XX  
 KW Hypersensitive response elicitor; fruit; vegetable; plant; desiccation;  
 KW postharvest disease.  
 XX  
 OS Erwinia amylovora.  
 XX  
 PN WO200180639-A2.  
 XX  
 PD 01-NOV-2001.  
 XX  
 PF 17-APR-2001; 2001WO-US012468.  
 XX  
 PR 19-APR-2000; 2000US-0198359P.  
 XX  
 PA (EDEN-) EDEN BIOSCIENCE CORP.  
 XX  
 PI Wei Z, Qiu D, Remick D;  
 XX  
 WI 2002-041357/05.  
 DR N-PSDB; ABL51711.  
 XX  
 PT Inhibiting post harvest disease (caused by Penicillium, Botrytis,  
 PT Phytophthora, or Erwinia) or desiccation and enhancing the longevity in a  
 PT fruits or vegetables, using hypersensitive response elicitor proteins or  
 PT nucleic acids.  
 XX  
 PS Disclosure; Page 11-12; 72pp; English.  
 XX  
 CC The present invention describes methods for inhibiting post harvest  
 CC disease or desiccation and enhancing the longevity in a fruits or  
 CC vegetables, using hypersensitive response elicitor proteins or  
 CC polypeptides or nucleic acids (i) derived from pathogens (e.g. Erwinia  
 CC amylovora, E. stewartii, E. chrysanthemi, E. carotovora, Xanthomonas,  
 CC Pseudomonas syringae, P. solanacearum, Phytophthora, and Clavibacter).  
 CC (i) has bactericidal activity, and can be used in gene therapy. The  
 CC method can be used for inhibiting post harvest disease (caused by  
 CC Penicillium, Botrytis, Phytophthora, or Erwinia) or desiccation and  
 CC enhancing the longevity in a fruits or vegetables. The method enables  
 CC growers, warehouse packers, shippers and suppliers to process, handle and  
 CC store fruit and vegetables with reduced losses caused by post harvest



CC disease and desiccation, therefore reducing costs to the consumer and  
CC improving quality. The present sequence represents a hypersensitive  
CC response elicitor protein given in the exemplification of the present  
CC invention  
XX  
SQ Sequence 447 AA;

Query Match 100.0%; Score 128; DB 5; Length 447;  
Best Local Similarity 100.0%; Pred. No. 6.4e-11;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NSALGQOPIDRTIQEQMAQLLAELLS 27  
Db 31 NSALGQOPIDRTIQEQMAQLLAELLS 57  
|||||

RESULT 8  
ABO73754  
ID ABO73754 standard; protein; 313 AA.  
XX  
AC ABO73754;  
XX  
DT 29-JUL-2004 (first entry)  
XX  
DE Pseudomonas aeruginosa polypeptide #5929.  
XX  
KW Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.  
XX  
OS Pseudomonas aeruginosa.  
XX  
PN US6551795-B1.  
XX  
PD 22-APR-2003.  
XX  
PF 18-FEB-1999; 99US-00252991.  
XX  
PR 18-FEB-1998; 98US-0074788P.  
XX  
PR 27-JUL-1998; 98US-0094190P.  
XX  
PA (GENO-) GENOME THERAPEUTICS CORP.  
XX  
PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;  
XX  
WPI; 2003-615309/58.  
XX  
DR N-PSDB; ABD07325.  
XX  
PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,  
XX useful as molecular targets for diagnostics, prophylaxis and treatment of  
XX pathological conditions resulting from bacterial infection.  
XX  
PS Disclosure; SEQ ID NO 22500; 455pp; English.  
XX  
CC The invention relates to Pseudomonas aeruginosa polypeptides and the  
XX polynucleotides encoding them. The sequences are useful in diagnosis and  
XX therapy of pathological conditions, as molecular targets for diagnostics,  
XX prophylaxis and treatment of pathological conditions resulting from a  
XX bacterial infection, for evaluating a compound, such as a polypeptide,  
XX for the ability to bind a P. aeruginosa nucleic acid, as components of  
XX effective antibacterial targets, as targets for antibacterial drugs,  
XX including anti-P. aeruginosa drugs, as templates for recombinant  
XX production of P. aeruginosa-derived peptides or polypeptides, as target  
XX components for diagnosis and/or treatment of P. aeruginosa-caused  
XX infection, and in detection of P. aeruginosa sequences or other sequences  
XX of Pseudomonas species using biochip technology. Sequences ABO67826-  
XX ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The  
XX sequence data for this patent did not form part of the printed  
XX specification but was obtained in electronic format from USPTO at  
XX seqdata.uspto.gov/sequence.html  
XX  
SQ Sequence 313 AA;

Query Match 40.6%; Score 52; DB 7; Length 313;  
Best Local Similarity 55.6%; Pred. No. 19;

Matches 10; Conservative 5; Mismatches 3; Indels 0; Gaps 0;  
Qy 7 QPIDRTIQEQMAQLLAEL 24  
Db 21 QRLERQWVAQVLAEL 38  
|||||

RESULT 9  
ABO73936  
ID ABO73936 standard; protein; 457 AA.  
XX  
AC ABO73936;  
XX  
DT 29-JUL-2004 (first entry)  
XX  
DE Pseudomonas aeruginosa polypeptide #6111.  
XX  
KW Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.  
XX  
OS Pseudomonas aeruginosa.  
XX  
PN US6551795-B1.  
XX  
PD 22-APR-2003.  
XX  
PF 18-FEB-1999; 99US-00252991.  
XX  
PR 18-FEB-1998; 98US-0074788P.  
XX  
PR 27-JUL-1998; 98US-0094190P.  
XX  
PA (GENO-) GENOME THERAPEUTICS CORP.  
XX  
PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;  
XX  
WPI; 2003-615309/58.  
XX  
DR N-PSDB; ABD07507.  
XX  
PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,  
XX useful as molecular targets for diagnostics, prophylaxis and treatment of  
XX pathological conditions resulting from bacterial infection.  
XX  
PS Disclosure; SEQ ID NO 22682; 455pp; English.  
XX  
CC The invention relates to Pseudomonas aeruginosa polypeptides and the  
XX polynucleotides encoding them. The sequences are useful in diagnosis and  
XX therapy of pathological conditions, as molecular targets for diagnostics,  
XX prophylaxis and treatment of pathological conditions resulting from a  
XX bacterial infection, for evaluating a compound, such as a polypeptide,  
XX for the ability to bind a P. aeruginosa nucleic acid, as components of  
XX effective antibacterial targets, as targets for antibacterial drugs,  
XX including anti-P. aeruginosa drugs, as templates for recombinant  
XX production of P. aeruginosa-derived peptides or polypeptides, as target  
XX components for diagnosis and/or treatment of P. aeruginosa-caused  
XX infection, and in detection of P. aeruginosa sequences or other sequences  
XX of Pseudomonas species using biochip technology. Sequences ABO67826-  
XX ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The  
XX sequence data for this patent did not form part of the printed  
XX specification but was obtained in electronic format from USPTO at  
XX seqdata.uspto.gov/sequence.html  
XX  
SQ Sequence 457 AA;

Query Match 40.6%; Score 52; DB 7; Length 457;  
Best Local Similarity 50.0%; Pred. No. 30;  
Matches 11; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Qy 4 LGQOPIDRTIQEQMAQLLAEL 25  
Db 312 LGQSPLLAALABELLAEL 333  
|||||

RESULT 10  
ADN73161











Qy 1 NSALGQ-QPIDRQTIEQMAQ 19  
| | | | | : | :  
Db 390 NSALGQTQPSDQDTLVQRAE 409

Search completed: January 25, 2005, 12:11:58  
Job time : 68.75 secs





GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2005, 12:08:41 ; Search time 30.6346 Seconds  
(without alignments)  
58.450 Million cell updates/sec

Title: US-09-879-248-6\_COPY\_31\_57  
Perfect score: 128  
Sequence: 1 NSALGOQPIDRQTIEQMAQLLALLKS 27

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 56318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 100 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/prodata/1/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/prodata/1/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/prodata/1/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/prodata/1/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/prodata/1/iaa/PCITUS\_COMB.pep.\*  
6: /cgn2\_6/prodata/1/iaa/backfile1.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	128	100.0	447	3	US-09-120-927-2
2	128	100.0	447	4	US-09-431-614-6
3	52	40.6	313	4	US-09-252-991A-22500
4	52	40.6	457	4	US-09-252-991A-22682
5	51	39.8	1312	2	US-08-592-126-148
6	51	39.8	1312	2	US-08-687-080-51
7	51	39.8	1312	4	US-09-168-595-148
8	49	38.3	173	4	US-09-270-767-40996
9	49	38.3	173	4	US-09-270-767-56212
10	49	38.3	264	4	US-09-252-991A-20342
11	48.5	37.9	831	4	US-09-252-991A-19010
12	48.5	37.9	891	4	US-09-252-991A-25595
13	47	36.7	112	4	US-09-489-039A-9733
14	47	36.7	505	4	US-09-252-991A-23615
15	45	35.2	620	3	US-08-982-785A-10
16	45	35.2	620	4	US-09-629-498-10
17	45	35.2	631	4	US-09-252-991A-20418
18	45	35.2	686	3	US-08-982-785A-8
19	45	35.2	686	4	US-09-629-498-8
20	45	35.2	756	3	US-08-982-785A-2
21	45	35.2	756	4	US-09-629-498-2
22	45	35.2	1084	3	US-09-227-725A-3
23	45	35.2	1225	4	US-09-501-171-4
24	45	35.2	1247	4	US-09-501-171-6
25	44.5	34.8	122	4	US-09-248-796A-28164
26	44	34.4	130	4	US-09-107-532A-4295
27	44	34.4	255	4	US-09-252-991A-31162

28	44	34.4	319	1	US-08-125-746-1	Sequence 1, Appli
29	44	34.4	319	4	US-08-948-276-2	Sequence 2, Appli
30	44	34.4	320	1	US-08-125-746-3	Sequence 3, Appli
31	44	34.4	320	4	US-08-948-276-1	Sequence 1, Appli
32	44	34.4	320	6	5225537-4	Patent No. 5225537
33	44	34.4	325	4	US-09-248-796A-20030	Sequence 20030, A
34	44	34.4	325	3	US-09-324-096A-2	Sequence 2, Appli
35	44	34.4	327	3	US-09-324-096A-4	Sequence 4, Appli
36	44	34.4	327	3	US-09-324-096A-6	Sequence 6, Appli
37	44	34.4	328	4	US-09-248-796A-24205	Sequence 24205, A
38	44	34.4	378	4	US-09-710-279-1428	Sequence 1428, Ap
39	44	34.4	387	3	US-09-134-001C-3248	Sequence 3248, Ap
40	44	34.4	457	4	US-08-924-629C-6	Sequence 6, Appli
41	44	34.4	502	4	US-10-101-464A-945	Sequence 945, App
42	44	34.4	635	4	US-09-252-991A-20298	Sequence 20298, A
43	43	33.6	93	4	US-09-489-039A-7880	Sequence 7880, Ap
44	43	33.6	190	4	US-09-252-991A-29854	Sequence 29854, A
45	43	33.6	315	4	US-09-252-991A-31381	Sequence 31381, A
46	43	33.6	430	4	US-08-956-171B-5244	Sequence 5244, Ap
47	43	33.6	430	4	US-08-781-986A-5244	Sequence 5244, Ap
48	43	33.6	883	4	US-09-543-681A-6612	Sequence 6612, Ap
49	43	33.6	983	3	US-09-134-001C-3814	Sequence 3814, Ap
50	43	33.6	1282	4	US-09-543-681A-5419	Sequence 5419, Ap
51	43	33.6	3072	3	US-09-413-814-93	Sequence 93, Appl
52	43	33.6	3079	3	US-09-413-814-80	Sequence 80, Appl
53	42.5	33.2	350	4	US-09-248-796A-16002	Sequence 16002, A
54	42	32.8	36	2	US-08-806-203-1	Sequence 1, Appli
55	42	32.8	161	4	US-09-107-532A-6339	Sequence 6339, Ap
56	42	32.8	184	4	US-09-248-796A-15181	Sequence 15181, A
57	42	32.8	191	4	US-09-270-767-44048	Sequence 44048, A
58	42	32.8	235	4	US-09-270-767-33215	Sequence 33215, A
59	42	32.8	238	4	US-09-270-767-31899	Sequence 31899, A
60	42	32.8	238	4	US-09-270-767-47116	Sequence 47116, A
61	42	32.8	254	4	US-09-252-991A-26502	Sequence 26502, A
62	42	32.8	328	4	US-09-270-767-59817	Sequence 59817, A
63	42	32.8	397	4	US-09-252-991A-18356	Sequence 18356, A
64	42	32.8	593	1	US-08-208-036-17	Sequence 17, Appl
65	42	32.8	593	1	US-08-428-823-17	Sequence 17, Appl
66	42	32.8	876	2	US-08-633-476-2	Sequence 2, Appli
67	42	32.8	877	1	US-08-208-036-14	Sequence 14, Appl
68	42	32.8	877	1	US-08-428-823-14	Sequence 14, Appl
69	42	32.8	984	4	US-09-252-991A-17368	Sequence 17368, A
70	42	32.8	1021	4	US-10-101-464A-954	Sequence 954, App
71	42	32.8	1388	4	US-09-270-767-44387	Sequence 44387, A
72	42	32.8	1548	1	US-08-463-092B-7	Sequence 7, Appli
73	42	32.8	1548	2	US-08-460-907B-7	Sequence 7, Appli
74	42	32.8	2004	1	US-08-375-709-15	Sequence 15, Appl
75	42	32.8	2004	1	US-08-752-929-15	Sequence 15, Appl
76	42	32.8	2004	3	US-09-090-793-9	Sequence 9, Appli
77	42	32.8	2004	4	US-09-231-899-9	Sequence 9, Appli
78	41.5	32.4	227	1	US-08-414-625-4	Sequence 4, Appli
79	41.5	32.4	367	4	US-09-248-796A-19641	Sequence 19641, A
80	41.5	32.4	448	4	US-09-543-681A-7737	Sequence 7737, Ap
81	41.5	32.4	539	4	US-09-252-991A-21493	Sequence 21493, A
82	41.5	32.4	582	4	US-09-248-796A-19158	Sequence 19158, A
83	41.5	32.4	2568	4	US-09-866-108A-3	Sequence 3, Appli
84	41	32.0	92	4	US-09-543-681A-7692	Sequence 7692, Ap
85	41	32.0	104	4	US-09-270-767-32777	Sequence 32777, Ap
86	41	32.0	104	4	US-09-270-767-47994	Sequence 47994, A
87	41	32.0	185	4	US-09-583-110-4159	Sequence 4159, Ap
88	41	32.0	302	4	US-09-134-000C-5055	Sequence 5055, Ap
89	41	32.0	352	4	US-09-252-991A-17455	Sequence 17455, A
90	41	32.0	372	4	US-09-248-796A-14919	Sequence 14919, A
91	41	32.0	453	4	US-09-248-796A-14922	Sequence 14922, A
92	41	32.0	484	1	US-08-597-236-2	Sequence 2, Appli
93	41	32.0	484	1	US-08-746-682A-2	Sequence 2, Appli
94	41	32.0	692	4	US-09-252-991A-26724	Sequence 26724, A
95	41	32.0	720	4	US-09-394-272-14	Sequence 14, Appl
96	41	32.0	1253	4	US-09-489-039A-12097	Sequence 12097, A
97	41	32.0	1428	4	US-09-252-991A-30731	Sequence 30731, A
98	41	32.0	1618	1	US-07-853-913-4	Sequence 4, Appli
99	41	32.0	1618	4	US-09-538-092-1143	Sequence 1143, Ap
100	41	32.0	2442	3	US-09-514-247A-10	Sequence 10, Appl



```

Qy 4 LGQSPIDRQTIEQMAQLLAELL 25
   |||:|
Db 312 LGQSPLLAALAEELLQLLAEML 333

RESULT 5
US-08-592-126-148
; Sequence 148, Application US/08592126
; Patent No. 5821091
; GENERAL INFORMATION:
; APPLICANT: Gregory Dolganov
; TITLE OF INVENTION: Transcripts Encoding Immunomodulatory
; TITLE OF INVENTION: Polypeptides
; NUMBER OF SEQUENCES: 151
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/592,126
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Sholtz, Charles K.
; REGISTRATION NUMBER: 38,615
; REFERENCE/DOCKET NUMBER: 4600-0111
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 148:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1312 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEetical: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: Rad50.pro-translation of SEQ ID NO:
US-08-592-126-148

Query Match 39.8%; Score 51; DB 2; Length 1312
Best Local Similarity 47.4%; Pred. No. 26;
Matches 9; Conservative 5; Mismatches 5; Indels

Qy 9 IDRQTIEQMAQLLAELLKS 27
   :||:|:|:|:|:|:|:|:|
Db 1237 LDRENIESLAHALVEIKS 1255

RESULT 6
US-08-687-080-51
; Sequence 51, Application US/08687080
; Patent No. 5965427
; GENERAL INFORMATION:
; APPLICANT: Gregory Dolganov
; TITLE OF INVENTION: Human RAD50 Gene and Methods of Use
; NUMBER OF SEQUENCES: 175
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306

```

```

; REFERENCE/DOCKET NUMBER: 4600-0111
;
; TELECOMMUNICATION INFORMATION:
;
; TELEPHONE: (415) 324-0880
;
; TELEFAX: (415) 324-0960
;
; INFORMATION FOR SEQ ID NO: 148:
;
; SEQUENCE CHARACTERISTICS:
;
; LENGTH: 1312 amino acids
;
; TYPE: amino acid
;
; TOPOLOGY: linear
;
; MOLECULE TYPE: protein
;
; HYPOTHETICAL: NO
;
; ORIGINAL SOURCE:
;
; INDIVIDUAL ISOLATE: Rad50.pro-tran
; US-09-168-595-148

```

Query Match 39.8%; Score 51; DB 4; Length 1312;  
Best Local Similarity 47.4%; Pred. No. 26;  
Matches 9; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

**Qy**            9 IDRTIEQMAQLLAELLS 27  
               :||: ||: | | |:|  
**Db**            1237 LDRENIESLAHALVEIKS 1255

```

RESULT 8
US-09-270-767-40996
; Sequence 40996, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and
; FILE REFERENCE: File Reference: 7326-0
; CURRENT APPLICATION NUMBER: US/09/270,
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 40996
; LENGTH: 173
; TYPE: prt
; ORGANISM: Drosophila melanogaster
US-09-270-767-40996

```

Query Match	38.3%	Score 49;	DB 4;	Length 173;
Best Local Similarity	44.4%;	Pred. No. 4.9;		
Matches 12;	Conservative	3;	Mismatches 12;	Indels 0;
Gaps 0;				

QY 1 NSALGQQPIDRQTIEQMAQLLAELLKS 27  
||| ||||| : ||| : ||| :  
Db 127 NSMLNQQPCSOOQLLATOLLYARLLRS 153

```

RESULT 9
US-09-270-767-56212
; Sequence 56212, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and
; FILE REFERENCE: File Reference: 7326-0
; CURRENT APPLICATION NUMBER: US/09270,
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 56212
; LENGTH: 173
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-56212

```

Query Match	38.3%	Score 49;	DB 4;	Length 173;
Best Local Similarity	44.4%;	Pred. No. 4.9;		
Matches 12;	Conservative	3;	Mismatches	12;
Indels	0;	Gaps	0;	

**Qy**                1 NSALGQQPIDRQTIEQAQLLAELKS 27  
                 ||| ||| : | : | | | : |  
**Dd**                127 NSMLNQPSCSQQLLATOLLYARLRS 153  
                 ||| ||| : | : | | | : |

RESULT 10  
US-09-252-991A-20342  
; Sequence 20342, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136

; PRIOR FILING DATE: 1998-07-27  
 ; PRIOR AFFILIATION NUMBER: US 6  
 ; NUMBER OF SEQ ID NOS: 33142  
 ; SEQ ID NO 20342

US-09-252-991A-20342

Query Match	38.3%	Score 49;	DB 4;	Length 264;
Best Local Similarity	58.8%;	Pred. No. 8;		
Matches 10;	Conservative	2;	Mismatches	5;
			Indels	0;
			Gaps	0;

Qy 3 ALGQQPIDRQTIEQMAQ 19  
||| | | | | | | | | |  
Db 196 AAGQHPDHRQETVEQEAQ 212

RESULT 11  
US-09-252-991A-19010  
; Sequence 19010, Application US/09252991A

```

: GENERAL INFORMATION: Rubenfield et al.
: APPLICANT: Marc J. NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
: TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
: TITLE OF INVENTION:

```

FILE REFERENCE: 10190.130  
CURRENT APPLICATION NUMBER: US/09/252,991A  
CURRENT FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/094,190  
 ; PRIOR FILING DATE: 1998-07-27  
 ; NUMBER OF SEQ ID NOS: 33142

```

; NUMBER OF SEQ ID NOS: 33142
;
; SEQ ID NO 19010
; LENGTH: 237
; TYPE: PRG
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19010

```

Query Match	37.9%;	Score 48.5;	DB 4;	Length 237;
Best Local Similarity	46.2%;	Pred. No. 8.4;		
Matches 12;	Conservative	6;	Mismatches	7;
			Indels	1;
			Gaps	1;

**Qy**            2 SALGQQ-PIDRQTTEOMAQLLAELLK 26  
             ||||| : | : | : | : | : | :  
**Dd**            26 SALPOEKPPLPROAFDDDPVLAEVLR 51

RESULT 12  
US-09-252-991A-25595  
; Sequence 25595, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.

```

; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-23615

Query Match          36.7%; Score 47; DB 4; Length 505;
Best Local Similarity 42.9%; Pred. No. 35;
Matches 12; Conservative 8; Mismatches 4; Indels

QY  3  ALGQQPIDRQTIEQM-----AQLLAELUK 26
      |||: |||: |||: |||: |||: |||:
DB  317  AVGQRPQDQRTVVQAGDADAQVVAFLLE 344

RESULT 15
US-08-982-785A-10
; Sequence 10, Application US/08982785A
; Patent No. 6258929
; GENERAL INFORMATION:
; APPLICANT: Kosik, Kenneth S.
; APPLICANT: Zhou, Jianhua
; TITLE OF INVENTION: ALARM RELATED PEPTIDES AND
; TITLE OF INVENTION: NUCLEIC ACIDS AND DIAGNOSIS USING THEM
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: FastSQ for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/982,785A
; FILING DATE: 02-DEC-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/031,556
; FILING DATE: 02-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Freeman, John W.
; REGISTRATION NUMBER: 29,066
; REFERENCE/DOCKET NUMBER: 05311/018001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 620 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-982-785A-10

Query Match          35.2%; Score 45; DB 3; Length 620;
Best Local Similarity 43.5%; Pred. No. 92;
Matches 10; Conservative 3; Mismatches 10; Indels

QY  2  SALGQQPIDRQTIEQMAQLLAEL 24
      |||| |||: |||: |||:
DB  315  SALGSSEIDSKVTENCVCILRNL 337

RESULT 16
US-09-629-498-10
; Sequence 10, Application US/09629498
; Patent No. 6797511
; GENERAL INFORMATION:
; APPLICANT: Kosik, Kenneth S.
; APPLICANT: Zhou, Jianhua
; TITLE OF INVENTION: ALARM RELATED PEPTIDES AND

```

NUCLEIC ACIDS AND DIAGNOSIS USING THEM

NUMBER OF SEQUENCES: 13

CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish & Richardson P.C.

STREET: 225 Franklin Street

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02110-2804

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: Windows 95

SOFTWARE: FastSEQ for Windows Version 2.0b

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/629,498

FILING DATE: 31-Jul-2000

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/982,785

FILING DATE: 02-DEC-1997

APPLICATION NUMBER: US 60/031,556

FILING DATE: 02-DEC-1996

ATTORNEY/AGENT INFORMATION:

NAME: Freeman, John W.

REGISTRATION NUMBER: 29,066

REFERENCE/DOCKET NUMBER: 05311/018001

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617/542-5070

TELEFAX: 617/542-8906

TELEX: 200154

INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:

LENGTH: 620 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 10:

US-09-629-498-10

Query Match 35.2%; Score 45; DB 4; Length 620;

Best Local Similarity 43.5%; Pred. No. 92;

Matches 10; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

Qy 2 SALGQQPIDQTIEQMAQLLAEL 24

Db 315 SALGSSSEIDSKTVENCVCILRNL 337

RESULT 17

US-09-252-991A-20418

Sequence 20418, Application US/09252991A

Patent No. 6551795

GENERAL INFORMATION:

APPLICANT: Marc J. Rubenfield et al.

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

FILE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: 107196.136

CURRENT APPLICATION NUMBER: US/09/252,991A

CURRENT FILING DATE: 1999-02-18

PRIOR APPLICATION NUMBER: US 60/074,788

PRIOR FILING DATE: 1998-02-18

PRIOR APPLICATION NUMBER: US 60/094,190

PRIOR FILING DATE: 1998-07-27

NUMBER OF SEQ ID NOS: 33142

SEQ ID NO 20418

LENGTH: 631

TYPE: PRT

ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-20418

Query Match 35.2%; Score 45; DB 4; Length 631;

Best Local Similarity 43.5%; Pred. No. 94;

Matches 10; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

Qy 3 ALGQQPIDQTIEQMAQLLAEL 25

Db 182 ALGEEHIGRTFLGQFLQALQLL 204

RESULT 18

US-08-982-785A-8

Sequence 8, Application US/08982785A

Patent No. 6258929

GENERAL INFORMATION:

APPLICANT: Kosik, Kenneth S.

APPLICANT: Zhou, Jianhua

TITLE OF INVENTION: ALARM RELATED PEPTIDES AND

TITLE OF INVENTION: NUCLEIC ACIDS AND DIAGNOSIS USING THEM

NUMBER OF SEQUENCES: 13

CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish & Richardson P.C.

STREET: 225 Franklin Street

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02110-2804

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: Windows 95

SOFTWARE: FastSEQ for Windows Version 2.0b

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/982,785A

FILING DATE: 02-DEC-1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 60/031,556

FILING DATE: 02-DEC-1996

ATTORNEY/AGENT INFORMATION:

NAME: Freeman, John W.

REGISTRATION NUMBER: 29,066

REFERENCE/DOCKET NUMBER: 05311/018001

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617/542-5070

TELEFAX: 617/542-8906

TELEX: 200154

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 686 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-982-785A-8

Query Match 35.2%; Score 45; DB 3; Length 686;

Best Local Similarity 43.5%; Pred. No. 1e+02;

Matches 10; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

Qy 2 SALGQQPIDQTIEQMAQLLAEL 24

Db 312 SALGSSSEIDSKTVENCVCILRNL 334

RESULT 19

US-09-629-498-8

Sequence 8, Application US/09629498

Patent No. 6797511

GENERAL INFORMATION:

APPLICANT: Kosik, Kenneth S.

APPLICANT: Zhou, Jianhua

TITLE OF INVENTION: ALARM RELATED PEPTIDES AND

TITLE OF INVENTION: NUCLEIC ACIDS AND DIAGNOSIS USING THEM

NUMBER OF SEQUENCES: 13

CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish & Richardson P.C.

STREET: 225 Franklin Street

CITY: Boston







GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2005, 12:09:52 ; Search time 50.3654 Seconds  
(without alignments)  
193.681 Million cell updates/sec

Title: US-09-879-248-6\_COPY\_31\_57

Perfect score: 128  
Sequence: 1 NSALGQPPDRQTIEQMAQLLALLKLS 27

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1608061 seqs, 361289386 residues

Total number of hits satisfying chosen parameters: 1608061

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : Published Applications AA.\*

1:	/cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2:	/cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3:	/cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4:	/cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5:	/cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6:	/cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7:	/cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8:	/cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9:	/cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10:	/cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11:	/cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12:	/cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13:	/cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14:	/cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15:	/cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16:	/cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
17:	/cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
18:	/cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
19:	/cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
20:	/cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	128	100.0	447	9	US-09-835-684-5
2	128	100.0	447	9	US-09-880-371-5
3	128	100.0	447	9	US-09-879-248-6
4	128	100.0	447	14	US-10-010-390-5
5	128	100.0	447	15	US-10-441-736-6
6	128	100.0	447	17	US-10-847-142-5
7	51	39.8	660	15	US-10-406-686A-50
8	51	39.8	1312	14	US-10-393-602-148
9	51	39.8	1318	16	US-10-408-765A-666
10	50.5	39.5	266	14	US-10-043-487-245
11	50.5	39.5	536	14	US-10-205-823-222
12	49	38.3	628	16	US-10-437-963-174471
13	49	38.3	697	14	US-10-183-687-244
					Sequence 5, Appli
					Sequence 5, Appli
					Sequence 6, Appli
					Sequence 5, Appli
					Sequence 5, Appli
					Sequence 5, Appli
					Sequence 50, Appl
					Sequence 148, App
					Sequence 666, App
					Sequence 245, App
					Sequence 222, App
					Sequence 174471,
					Sequence 244, App

14	49	38.3	3249	15	US-10-282-122A-67572	Sequence 67572, A
15	48.5	37.9	211	17	US-10-425-115-32898	Sequence 32898,
16	48	37.5	96	17	US-10-425-115-267186	Sequence 267186,
17	48	37.5	470	14	US-10-369-493-4813	Sequence 4813, Ap
18	48	37.5	478	14	US-10-369-493-7572	Sequence 7572, Ap
19	48	37.5	492	14	US-10-032-189-10	Sequence 10, Appl
20	48	37.5	495	14	US-10-032-189-68	Sequence 68, Appl
21	48	37.5	630	15	US-10-282-122A-49449	Sequence 49449, A
22	48	37.5	1419	16	US-10-473-115-2	Sequence 2, Appli
23	48	37.5	1449	16	US-10-408-765A-1810	Sequence 1810, Ap
24	48	37.5	2246	15	US-10-282-122A-69540	Sequence 69540, A
25	47	36.7	83	15	US-10-424-599-147641	Sequence 147641,
26	46	35.9	31	16	US-10-745-069-52	Sequence 52, Appl
27	46	35.9	31	17	US-10-768-288A-52	Sequence 52, Appl
28	46	35.9	31	17	US-10-780-325A-52	Sequence 52, Appl
29	46	35.9	31	17	US-10-869-649-52	Sequence 52, Appl
30	46	35.9	32	16	US-10-745-069-51	Sequence 51, Appl
31	46	35.9	32	17	US-10-768-288A-51	Sequence 51, Appl
32	46	35.9	32	17	US-10-780-325A-51	Sequence 51, Appl
33	46	35.9	32	17	US-10-869-649-51	Sequence 51, Appl
34	46	35.9	33	16	US-10-745-069-50	Sequence 50, Appl
35	46	35.9	33	17	US-10-768-288A-50	Sequence 50, Appl
36	46	35.9	33	17	US-10-780-325A-50	Sequence 50, Appl
37	46	35.9	33	17	US-10-869-649-50	Sequence 50, Appl
38	46	35.9	34	16	US-10-745-069-49	Sequence 49, Appl
39	46	35.9	34	17	US-10-768-288A-49	Sequence 49, Appl
40	46	35.9	34	17	US-10-780-325A-49	Sequence 49, Appl
41	46	35.9	34	17	US-10-869-649-49	Sequence 49, Appl
42	46	35.9	36	16	US-10-745-069-47	Sequence 47, Appl
43	46	35.9	36	17	US-10-768-288A-47	Sequence 47, Appl
44	46	35.9	36	17	US-10-780-325A-47	Sequence 47, Appl
45	46	35.9	36	17	US-10-869-649-47	Sequence 47, Appl
46	46	35.9	38	15	US-10-424-599-262864	Sequence 262864,
47	46	35.9	238	15	US-10-425-114-49572	Sequence 49572, A
48	46	35.9	259	15	US-10-425-114-49036	Sequence 49036, A
49	46	35.9	284	16	US-10-437-963-192612	Sequence 192612,
50	46	35.9	348	14	US-10-156-761-9234	Sequence 9234, Ap
51	46	35.9	449	9	US-09-738-626-6150	Sequence 6150, Ap
52	46	35.9	482	15	US-10-425-114-43041	Sequence 43041, A
53	46	35.9	497	14	US-10-369-493-9870	Sequence 9870, Ap
54	46	35.9	566	15	US-10-425-114-56072	Sequence 56072, A
55	46	35.9	863	15	US-10-424-599-155432	Sequence 155432, A
56	46	35.9	1010	15	US-10-424-599-155812	Sequence 155812,
57	46	35.9	1013	15	US-10-424-599-277819	Sequence 277819,
58	45.5	35.5	249	15	US-10-282-122A-65161	Sequence 65161, A
59	45.5	35.5	249	15	US-10-282-122A-65658	Sequence 65658, A
60	45.5	35.5	351	14	US-10-156-761-8050	Sequence 8050, Ap
61	45.5	35.5	422	15	US-10-282-122A-51218	Sequence 51218, A
62	45.5	35.5	620	15	US-10-425-114-59476	Sequence 59476, A
63	45	35.2	29	16	US-10-745-069-54	Sequence 54, Appl
64	45	35.2	29	17	US-10-768-288A-54	Sequence 54, Appl
65	45	35.2	29	17	US-10-780-325A-54	Sequence 54, Appl
66	45	35.2	29	17	US-10-869-649-54	Sequence 54, Appl
67	45	35.2	30	16	US-10-745-069-53	Sequence 53, Appl
68	45	35.2	30	17	US-10-768-288A-53	Sequence 53, Appl
69	45	35.2	30	17	US-10-780-325A-53	Sequence 53, Appl
70	45	35.2	30	17	US-10-869-649-53	Sequence 53, Appl
71	45	35.2	36	16	US-10-745-069-85	Sequence 85, Appl
72	45	35.2	36	17	US-10-768-288A-85	Sequence 85, Appl
73	45	35.2	36	17	US-10-780-325A-85	Sequence 85, Appl
74	45	35.2	36	17	US-10-869-649-85	Sequence 85, Appl
75	45	35.2	50	17	US-10-776-013-614	Sequence 614, App
76	45	35.2	50	17	US-10-776-013-614	Sequence 614, App
77	45	35.2	95	9	US-09-757-713-2	Sequence 2, Appli
78	45	35.2	128	15	US-10-424-599-173202	Sequence 173202,
79	45	35.2	134	17	US-10-425-115-282746	Sequence 282746,
80	45	35.2	136	15	US-10-425-114-41859	Sequence 41859, A
81	45	35.2	136	15	US-10-425-114-45859	Sequence 45859, A
82	45	35.2	147	16	US-10-767-701-40481	Sequence 40481, A
83	45	35.2	162	15	US-10-424-599-171006	Sequence 171006,
84	45	35.2	252	9	US-09-738-626-5093	Sequence 5093, Ap
85	45	35.2	307	14	US-10-190-471-4	Sequence 4, Appli
86	45	35.2	489	9	US-09-815-242-5850	Sequence 5850, Ap

87 45 35.2 561 17 US-10-425-115-282743 Sequence 282743,  
88 45 35.2 705 10 US-09-907-907A-43 Sequence 43, Appl  
89 45 35.2 872 14 US-10-369-493-8935 Sequence 8935, Ap  
90 45 35.2 978 15 US-10-282-122A-44151 Sequence 44151, A  
91 45 35.2 981 9 US-09-815-242-12211 Sequence 12211, A  
92 45 35.2 1029 16 US-10-437-963-164029 Sequence 164029,  
93 45 35.2 1084 13 US-10-071-900-3 Sequence 3, Appl  
94 45 35.2 1818 15 US-10-282-122A-64245 Sequence 64245, A  
95 45 35.2 2020 16 US-10-437-963-195615 Sequence 195615,  
96 45 35.2 3390 14 US-10-369-493-21101 Sequence 21101, A  
97 45 35.2 11300 16 US-10-250-304A-2 Sequence 2, Appl  
98 44.5 34.8 739 17 US-10-425-115-211368 Sequence 211368,  
99 44.5 34.8 1153 15 US-10-282-122A-61350 Sequence 61350, A  
100 44 34.4 36 16 US-10-745-069-86 Sequence 86, Appl

## ALIGNMENTS

RESULT 1  
US-09-835-684-5  
; Sequence 5, Application US/09835684  
; Patent No. US20020019337A1  
; GENERAL INFORMATION:  
; APPLICANT: Wei, Zhong-Min  
; APPLICANT: Qiu, Dewen  
; APPLICANT: Remick, Dean  
; TITLE OF INVENTION: TREATMENT OF FRUITS OR VEGETABLES WITH HYPERSENSITIVE  
; TITLE OF INVENTION: RESPONSE ELICITOR TO CONTROL POSTHARVEST DISEASE OR  
; TITLE OF INVENTION: DESICCATION  
; FILE REFERENCE: 21829/71  
; CURRENT APPLICATION NUMBER: US/09/835,684  
; PRIOR FILING DATE: 2001-04-16  
; PRIOR FILING DATE: 60/198,359  
; PRIOR FILING DATE: 2000-04-19  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 5  
; LENGTH: 447  
; TYPE: PRT  
; ORGANISM: Erwinia amylovora  
US-09-835-684-5

Query Match 100.0%; Score 128; DB 9; Length 447;  
Best Local Similarity 100.0%; Pred. No. 2.1e-10;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSALGQPPIDRQTIEQMAQLLAELLKS 27  
|||||

Db 31 NSALGQPPIDRQTIEQMAQLLAELLKS 57  
|||||

RESULT 2  
US-09-880-371-5  
; Sequence 5, Application US/09880371  
; Patent No. US20020059658A1  
; GENERAL INFORMATION:  
; APPLICANT: Wei, Zhong-Min  
; APPLICANT: Derocher, Jay  
; TITLE OF INVENTION: METHODS OF IMPROVING THE EFFECTIVENESS OF TRANSGENIC  
; TITLE OF INVENTION: PLANTS  
; FILE REFERENCE: 21829/91  
; CURRENT APPLICATION NUMBER: US/09/880,371  
; CURRENT FILING DATE: 2001-06-13  
; PRIOR FILING DATE: 60/211,585  
; PRIOR FILING DATE: 2000-06-15  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 5  
; LENGTH: 447  
; TYPE: PRT  
; ORGANISM: Erwinia amylovora  
US-09-880-371-5

Query Match 100.0%; Score 128; DB 9; Length 447;  
Best Local Similarity 100.0%; Pred. No. 2.1e-10;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSALGQPPIDRQTIEQMAQLLAELLKS 27  
|||||

Db 31 NSALGQPPIDRQTIEQMAQLLAELLKS 57  
|||||

RESULT 3  
US-09-879-248-6  
; Sequence 6, Application US/09879248  
; Patent No. US20020062500A1  
; GENERAL INFORMATION:  
; APPLICANT: Fan, Hao  
; APPLICANT: Wei, Zhong-Min  
; TITLE OF INVENTION: HYPERSENSITIVE RESPONSE ELICITING DOMAINS AND USE  
; TITLE OF INVENTION: THEREOF  
; FILE REFERENCE: 21829/81  
; CURRENT APPLICATION NUMBER: US/09/879,248  
; PRIOR FILING DATE: 2001-06-12  
; PRIOR FILING DATE: 60/212,211  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 6  
; LENGTH: 447  
; TYPE: PRT  
; ORGANISM: Erwinia amylovora  
US-09-879-248-6

Query Match 100.0%; Score 128; DB 9; Length 447;  
Best Local Similarity 100.0%; Pred. No. 2.1e-10;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSALGQPPIDRQTIEQMAQLLAELLKS 27  
|||||

Db 31 NSALGQPPIDRQTIEQMAQLLAELLKS 57  
|||||

RESULT 4  
US-10-010-390-5  
; Sequence 5, Application US/10010390  
; Publication No. US20030104979A1  
; GENERAL INFORMATION:  
; APPLICANT: Wei, Zhong-Min  
; APPLICANT: Leon, Ernesto  
; APPLICANT: Oviedo, Agustín  
; TITLE OF INVENTION: METHODS OF INHIBITING DESICCATION OF CUTTINGS REMOVED  
; TITLE OF INVENTION: FROM ORNAMENTAL PLANTS  
; FILE REFERENCE: 21829/111  
; CURRENT APPLICATION NUMBER: US/10/010,390  
; CURRENT FILING DATE: 2001-11-05  
; PRIOR FILING DATE: 60/248,169  
; PRIOR FILING DATE: 2000-11-13  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 5  
; LENGTH: 447  
; TYPE: PRT  
; ORGANISM: Erwinia amylovora  
US-10-010-390-5

Query Match 100.0%; Score 128; DB 14; Length 447;  
Best Local Similarity 100.0%; Pred. No. 2.1e-10;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSALGQPPIDRQTIEQMAQLLAELLKS 27  
|||||

Db 31 NSALGQPPIDRQTIEQMAQLLAELLKS 57  
|||||

```
RESULT 5
US-10-441-736-6
; Sequence 6, Application US/10441736
; Publication No. US20040016029A1
; GENERAL INFORMATION:
; APPLICANT: Wei, Zhong-Min
; APPLICANT: Schading, Richard L.
; TITLE OF INVENTION: HYPERSENSITIVE RESPONSE ELICITOR-INDUCED STRESS
; TITLE OF INVENTION: RESISTANCE
; FILE REFERENCE: 21829/203 (EBC-003)
; CURRENT APPLICATION NUMBER: US/10/441,736
; CURRENT FILING DATE: 2003-05-20
; PRIOR APPLICATION NUMBER: 60/107,243
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: 09/431,614
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 447
; TYPE: PRT
; ORGANISM: Erwinia amylovora
US-10-441-736-6

Query Match      100.0%; Score 128; DB 15; Length 447;
Best Local Similarity 100.0%; Pred. No. 2.1e-10;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSALGQOPIDRQTIEQMAQLLAELLKS 27
Db 31 NSALGQOPIDRQTIEQMAQLLAELLKS 57

RESULT 6
US-10-847-142-5
; Sequence 5, Application US/10847142
; Publication No. US20040265442A1
; GENERAL INFORMATION:
; APPLICANT: Wei, Zhong-Min
; APPLICANT: Qiu, Dewen
; APPLICANT: Remick, Dean
; TITLE OF INVENTION: TREATMENT OF FRUITS OR VEGETABLES WITH HYPERSENSITIVE
; TITLE OF INVENTION: RESPONSE ELICITOR TO CONTROL POSTHARVEST DISEASE OR
; TITLE OF INVENTION: DESICCATION
; FILE REFERENCE: 21829/197
; CURRENT APPLICATION NUMBER: US/10/847,142
; CURRENT FILING DATE: 2004-05-17
; PRIOR APPLICATION NUMBER: 60/198,359
; PRIOR FILING DATE: 2000-04-19
; PRIOR APPLICATION NUMBER: 09/835,684
; PRIOR FILING DATE: 2001-04-16
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 447
; TYPE: PRT
; ORGANISM: Erwinia amylovora
US-10-847-142-5

Query Match      100.0%; Score 128; DB 17; Length 447;
Best Local Similarity 100.0%; Pred. No. 2.1e-10;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSALGQOPIDRQTIEQMAQLLAELLKS 27
Db 31 NSALGQOPIDRQTIEQMAQLLAELLKS 57

RESULT 7
US-10-406-686A-50
; Sequence 50, Application US/10406686A
; Publication No. US20040033586A1
; GENERAL INFORMATION:
```

```
; APPLICANT: CROOKE, HELEN RACHEL
; APPLICANT: SHEA, JACQUELINE ELIZABETH
; APPLICANT: FELDMAN, ROBERT GRAHAM
; APPLICANT: COUTEBROZE, SYLVAIN GABRIEL
; APPLICANT: LEGROS, FRANCOIS-XAVIER
; TITLE OF INVENTION: ATTENUATED GRAM NEGATIVE BACTERIA
; FILE REFERENCE: 454313-3171.1
; CURRENT APPLICATION NUMBER: US/10/406,686A
; CURRENT FILING DATE: 2003-04-03
; PRIOR APPLICATION NUMBER: 60/370,282
; PRIOR FILING DATE: 2002-04-05
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 50
; LENGTH: 660
; TYPE: PRT
; ORGANISM: Pasteurella multocida
US-10-406-686A-50

Query Match      39.8%; Score 51; DB 15; Length 660;
Best Local Similarity 55.0%; Pred. No. 83;
Matches 11; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 6 QQPIDRQTIEQMAQLLAELL 25
Db 178 QQPHNKQKIEQLKQOELL 197

RESULT 8
US-10-393-602-148
; Sequence 148, Application US/10393602
; Publication No. US20030170714A1
; GENERAL INFORMATION:
; APPLICANT: Gregory Dolganov
; TITLE OF INVENTION: Transcripts Encoding Immunomodulatory
; NUMBER OF SEQUENCES: 151
; CORRESPONDENCE ADDRESS: Polypeptides
; ADDRESS: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/393,602
; FILING DATE: 19-Mar-2003
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/592,126
; FILING DATE: 26-JAN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Sholtz, Charles K.
; REGISTRATION NUMBER: 38,615
; REFERENCE/DOCKET NUMBER: 4600-0111
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 148:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1312 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: Rad50.pro-translation of SEQ ID NO:54
; SEQUENCE DESCRIPTION: SEQ ID NO: 148:
```

;; CURRENT FILING DATE: 2003-05-14  
;; NUMBER OF SEQ ID NOS: 204966  
;; SEQ ID NO 174471  
;; LENGTH: 628

```
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_72408C.1.pep
US-10-437-963-174471

Query Match      38.3%; Score 49; DB 16; Length 628;
Best Local Similarity 43.3%; Pred. No. 1.6e+02;
Matches 13; Conservative 4; Mismatches 9; Indels 4; Gaps 1;

QY      1 NSALGQGPIDRQT-----IEQMAQLLAELLK 26
      | : | | | | | | | | | | | | | | | | |
Db      415 NDAMDQPPSDCLTRVRLLEQVASTIAELVK 444

RESULT 13
US-10-183-687-244
; Sequence 244, Application US/10183687
; Publication No. US20030204870A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Allen, William B.
; APPLICANT: Cahoon, Rebecca
; APPLICANT: Epeibaum, Sabine
; APPLICANT: Famodu, Omolayo O.
; APPLICANT: Harvell, Leslie T.
; APPLICANT: Jones, Todd
; APPLICANT: Kinney, Tony
; APPLICANT: Klein, Ted
; APPLICANT: Li, Changjiang
; APPLICANT: Oliveira, Igor Cunha
; APPLICANT: Sakai, Hajime
; APPLICANT: Shen, Bo
; APPLICANT: Tarczynski, Mitchell C.
; TITLE OF INVENTION: Alteration Of Oil Traits In Plants
; FILE REFERENCE: BRL458 US NA
; CURRENT APPLICATION NUMBER: US/10/183,687
; CURRENT FILING DATE: 2002-06-27
; PRIOR APPLICATION NUMBER: 60/301,913
; PRIOR FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 532
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 244
; LENGTH: 697
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-183-687-244

Query Match      38.3%; Score 49; DB 14; Length 697;
Best Local Similarity 43.3%; Pred. No. 1.8e+02;
Matches 13; Conservative 4; Mismatches 9; Indels 4; Gaps 1;

QY      1 NSALGQGPIDRQT-----IEQMAQLLAELLK 26
      | : | | | | | | | | | | | | | | | | |
Db      484 NDAMDQPPSDCLTRVRLLEQVASTIAELVK 513

RESULT 14
US-10-282-122A-67572
; Sequence 67572, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.

; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 67572
; LENGTH: 3249
; TYPE: PRT
; ORGANISM: Pseudomonas putida
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (650)..(650)
; OTHER INFORMATION: X=any amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (1081)..(1081)
; OTHER INFORMATION: X=any amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (1088)..(1088)
; OTHER INFORMATION: X=any amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (1114)..(1114)
; OTHER INFORMATION: X=any amino acid
US-10-282-122A-67572

Query Match      38.3%; Score 49; DB 15; Length 3249;
Best Local Similarity 61.1%; Pred. No. 1e+03;
Matches 11; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY      10 DROTIEQMAQLLAELLK 27
      | | | | | | | | | | | | | | | | |
Db      2637 DROTIEQMAHWNLLQA 2654

RESULT 15
US-10-425-115-328898
; Sequence 328898, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
```

; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 328898  
; LENGTH: 211  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(211)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_63023C.1.pbp  
US-10-425-115-328898

Query Match 37.9%; Score 48.5; DB 17; Length 211;  
Best Local Similarity 39.4%; Pred. No. 52;  
Matches 13; Conservative 4; Mismatches 3; Indels 13; Gaps 1;

QY 3 ALGOQPIDRQTEQ-----MAQLLA 22  
Db 156 ALQMPIDRVGLEKEATALSRMGLRLTAQLA 188

RESULT 16  
US-10-425-115-267186  
; Sequence 267186, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; TITLE OF INVENTION: Plants  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 267186  
; LENGTH: 96  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_175277C.1.pbp  
US-10-425-115-267186

Query Match 37.5%; Score 48; DB 17; Length 96;  
Best Local Similarity 47.4%; Pred. No. 25;  
Matches 9; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 3 ALGOQPIDRQTEQMAQLL 21  
Db 8 SLGEHPLDKSTILQVAHQL 26

RESULT 17  
US-10-369-493-4813  
; Sequence 4813, Application US/10369493  
; Publication No. US20030233675A1  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Chen, Xianfeng  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES  
; FILE REFERENCE: 38-10(52052)B  
; CURRENT APPLICATION NUMBER: US/10/369,493  
; CURRENT FILING DATE: 2003-02-28  
; PRIOR APPLICATION NUMBER: US 60/360,039  
; PRIOR FILING DATE: 2002-02-21  
; NUMBER OF SEQ ID NOS: 47374  
; SEQ ID NO 4813

; LENGTH: 470  
; TYPE: PRT  
; ORGANISM: Burkholderia fungorum  
US-10-369-493-4813  
Query Match 37.5%; Score 48; DB 14; Length 470;  
Best Local Similarity 50.0%; Pred. No. 1.6e+02;  
Matches 13; Conservative 3; Mismatches 8; Indels 2; Gaps 1;  
QY 2 SALGOQPIDRQTEQMAQLLAELLKS 27  
Db 315 SAAGKYPV--QTETMAAICIEAEKS 338

RESULT 18  
US-10-369-493-7572  
; Sequence 7572, Application US/10369493  
; Publication No. US20030233675A1  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Chen, Xianfeng  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES  
; FILE REFERENCE: 38-10(52052)B  
; CURRENT APPLICATION NUMBER: US/10/369,493  
; CURRENT FILING DATE: 2003-02-28  
; PRIOR APPLICATION NUMBER: US 60/360,039  
; PRIOR FILING DATE: 2002-02-21  
; NUMBER OF SEQ ID NOS: 47374  
; SEQ ID NO 7572  
; LENGTH: 478  
; TYPE: PRT  
; ORGANISM: Burkholderia cepacia  
US-10-369-493-7572

Query Match 37.5%; Score 48; DB 14; Length 478;  
Best Local Similarity 50.0%; Pred. No. 1.6e+02;  
Matches 13; Conservative 3; Mismatches 8; Indels 2; Gaps 1;

QY 2 SALGOQPIDRQTEQMAQLLAELLKS 27  
Db 321 SAAGKYPV--QTETMAAICIEAEKS 344

RESULT 19  
US-10-032-189-10  
; Sequence 10, Application US/10032189  
; Publication No. US20030170630A1  
; GENERAL INFORMATION:  
; APPLICANT: Alsobrook II, John P  
; APPLICANT: Tchernev, Velizar T  
; APPLICANT: Liu, Xiaohong  
; APPLICANT: Spytek, Kimberly A  
; APPLICANT: Zerhusen, Bryan D  
; APPLICANT: Patturajan, Meera  
; APPLICANT: Grosse, William M  
; APPLICANT: Lepley, Denise M  
; APPLICANT: Burgess, Catherine E  
; APPLICANT: Shimkets, Richard A  
; APPLICANT: Grosse, William M  
; APPLICANT: Szekeres, Edward S  
; APPLICANT: Vernet, Corine A.M.  
; APPLICANT: Li, Li  
; APPLICANT: Casman, Stacie J  
; APPLICANT: Boldog, Ferenc L  
; APPLICANT: Gorman, Linda  
; APPLICANT: Gangolli, Esha A  
; APPLICANT: Fernandes, Elma R  
; APPLICANT: Rieger, Daniel K  
; APPLICANT: Edinger, Shlomit R

```

; APPLICANT: Gunther, Erik
; APPLICANT: Millet, Isabelle
; APPLICANT: Sciore, Paul
; APPLICANT: Ellerman, Karen
; APPLICANT: MacDougall, John R
; APPLICANT: Smithson, Glennda
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-228
; CURRENT APPLICATION NUMBER: US/10/032,189
; CURRENT FILING DATE: 2001-12-21
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/257,495
; PRIOR FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: 60/258,171
; PRIOR FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: 60/269,940
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/274,192
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/277,826
; PRIOR FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: 60/279,840
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/282,981
; PRIOR FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 60/283,656
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/309,247
; PRIOR FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: 60/311,754
; PRIOR FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: 60/313,331
; PRIOR FILING DATE: 2001-08-17
; NUMBER OF SEQ ID NOS: 260
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 492
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (492)
; OTHER INFORMATION: Wherein Xaa is any amino acid.
US-10-032-189-10

Query Match          37.5%; Score 48; DB 14; Length 492;
Best Local Similarity 37.0%; Pred. No. 1.6e+02;
Matches 10; Conservative 9; Mismatches 8; Indels 0; Gaps 0;

QY      1 NSALGQOPIRQTIQMAQLLAELLS 27
DB      134 NFGMGKRSIEIRLIERGSEFLAELRKT 160

RESULT 20
US-10-032-189-68
; Sequence 68, Application US/10032189
; Publication No. US20030170630A1
; GENERAL INFORMATION:
; APPLICANT: Alcobrook II, John P
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Liu, Xiaohong
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Zerhusen, Bryan D
; APPLICANT: Patturajan, Meera
; APPLICANT: Grosse, William M
; APPLICANT: Lepley, Denise M
; APPLICANT: Burgess, Catherine E
; APPLICANT: Shimkets, Richard A
; APPLICANT: Grosse, William M
; APPLICANT: Szekeres, Edward S
; APPLICANT: Vernet, Corine A.M.
; APPLICANT: Li, Li
; APPLICANT: Casman, Stacie J

```

```

; APPLICANT: Boldog, Ferenc L
; APPLICANT: Gorman, Linda
; APPLICANT: Gangolli, Bsha A
; APPLICANT: Fernandes, Elma R
; APPLICANT: Rieger, Daniel K
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Gunther, Erik
; APPLICANT: Millet, Isabelle
; APPLICANT: Sciore, Paul
; APPLICANT: Ellerman, Karen
; APPLICANT: MacDougall, John R
; APPLICANT: Smithson, Glennda
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-228
; CURRENT APPLICATION NUMBER: US/10/032,189
; CURRENT FILING DATE: 2001-12-21
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/257,495
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/258,171
; PRIOR FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: 60/269,940
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/274,192
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/277,826
; PRIOR FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: 60/279,840
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/282,981
; PRIOR FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 60/283,656
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/309,247
; PRIOR FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: 60/311,754
; PRIOR FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: 60/313,331
; PRIOR FILING DATE: 2001-08-17
; NUMBER OF SEQ ID NOS: 260
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 68
; LENGTH: 495
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (216)..(226)
; OTHER INFORMATION: Wherein Xaa is any amino acid.
US-10-032-189-68

Query Match          37.5%; Score 48; DB 14; Length 495;
Best Local Similarity 37.0%; Pred. No. 1.7e+02;
Matches 10; Conservative 9; Mismatches 8; Indels 0; Gaps 0;

QY      1 NSALGQOPIRQTIQMAQLLAELLS 27
DB      134 NFGMGKRSIEIRLIERGSEFLAELRKT 160

Search completed: January 25, 2005, 12:18:53
Job time : 53.3654 secs

```





GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2005, 12:08:06 ; Search time 11.9423 Seconds  
(without alignments)  
217.534 Million cell updates/sec

Title: US-09-879-248-6\_COPY\_31\_57

Perfect score: 128

Sequence: 1 NSALGQPIDRQTIEQMAQLLELLKS 27

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

1: PIR 79:\*

2: pir1:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	128	100.0	447	T18447	Hrpw protein - Erw
2	51	39.8	580	G69656	2-succinyl-6-hydro
3	51	39.8	1312	T30845	probable DNA repair
4	50	39.1	430	AE1803	sugar binding prot
5	50	39.1	430	AF1429	sugar binding prot
6	50	39.1	2706	T28155	variant-specific s
7	49	38.3	257	S76537	hypothetical prote
8	49	38.3	902	AH2999	conserved hypothet
9	49	38.3	903	A96284	hypothetical prote
10	48.5	37.9	213	C82953	hypothetical prote
11	48.5	37.9	627	S76462	hypothetical prote
12	48.5	37.9	762	H83348	probable acylase p
13	48	37.5	350	S76212	hypothetical prote
14	47	36.7	246	D87487	uridylylate kinase I
15	47	36.7	356	S01992	glutenin low molec
16	47	36.7	817	T01866	hypothetical prote
17	47	36.7	1122	T28130	hypothetical prote
18	46	35.9	36	PCBO	pancreatic hormone
19	46	35.9	59	1 PCSB	pancreatic hormone
20	46	35.9	75	G91013	hypothetical prote
21	46	35.9	75	A85858	hypothetical prote
22	46	35.9	75	B64988	hypothetical prote
23	46	35.9	159	AE2720	conserved hypothet
24	46	35.9	159	A97502	hypothetical prote
25	46	35.9	409	S76119	hypothetical prote
26	46	35.9	792	T29187	hypothetical prote
27	45.5	35.5	249	A81001	electron transfer
28	45.5	35.5	249	S75098	electron transfer
29	45.5	35.5	929	S75098	hypothetical prote

30	45	35.2	36	1	A61132	pancreatic hormone
31	45	35.2	36	1	C61132	pancreatic hormone
32	45	35.2	36	1	D61132	pancreatic hormone
33	45	35.2	66	1	PCCT	Mut7/nudix family
34	45	35.2	203	2	C95194	mutator protein li
35	45	35.2	203	2	H98060	conserved hypothet
36	45	35.2	455	2	B72339	hypothetical prote
37	45	35.2	510	2	AG2985	sugar ABC transpor
38	45	35.2	511	2	H98297	hypothetical prote
39	45	35.2	595	2	G96491	conserved hypothet
40	45	35.2	605	2	A82381	polyribonucleotide
41	45	35.2	705	2	S70691	conserved hypothet
42	45	35.2	978	2	B89971	neural plakophilin
43	45	35.2	1247	2	T42209	hypothetical prote
44	45	35.2	1818	1	S73852	COP9 signalosome c
45	44.5	34.8	397	2	T52302	probable portal pr
46	44.5	34.8	454	2	F90875	probable portal pr
47	44.5	34.8	833	2	H90821	unknown protein en
48	44.5	34.8	859	2	F85680	probable portal pr
49	44.5	34.8	859	2	H90909	probable portal pr
50	44.5	34.8	859	2	C90853	probable portal pr
51	44.5	34.8	861	2	B85844	hypothetical prote
52	44	34.4	58	2	S73287	hypothetical prote
53	44	34.4	201	2	B70616	hypothetical prote
54	44	34.4	209	2	A83524	hypothetical prote
55	44	34.4	216	2	D98240	probable hydrolase
56	44	34.4	226	2	F83307	probable hydrolase
57	44	34.4	229	2	B95287	probable hydrolase
58	44	34.4	229	2	AH3045	hydrolase (impor
59	44	34.4	262	2	S72841	hypothetical prote
60	44	34.4	262	2	F87071	conserved hypothet
61	44	34.4	299	2	A99497	heterodisulfide re
62	44	34.4	320	1	AQHUP	annexin V - bovine
63	44	34.4	320	1	S27214	annexin V - bovine
64	44	34.4	330	2	B71163	probable oligopept
65	44	34.4	381	2	D89922	conserved hypothet
66	44	34.4	417	2	A72236	hypothetical prote
67	44	34.4	503	2	A82193	Sun/nucleolar prot
68	44	34.4	597	2	B69251	probable electron
69	44	34.4	789	2	S62172	SHE4 protein - yea
70	44	34.4	896	2	S76064	hypothetical prote
71	44	34.4	896	2	S59990	phycobilisome anch
72	44	34.4	1002	2	T46033	receptor protein k
73	44	34.4	1003	2	T05898	hypothetical prote
74	44	34.4	1058	2	T19282	hypothetical prote
75	44	34.4	1069	2	D81826	exodeoxyribonuclea
76	43.5	34.0	449	2	D75033	probable prolifera
77	43.5	34.0	1499	2	AC2555	hypothetical prote
78	43	33.6	36	1	PCFG	pancreatic hormone
79	43	33.6	36	2	B60413	pancreatic hormone
80	43	33.6	93	1	PCDG	pancreatic hormone
81	43	33.6	131	2	T49331	hypothetical prote
82	43	33.6	179	2	B69406	conserved hypothet
83	43	33.6	182	2	B83539	hypothetical prote
84	43	33.6	209	2	T34530	hypothetical prote
85	43	33.6	212	2	AG0854	probable sugar ald
86	43	33.6	324	2	AE3075	transcription regu
87	43	33.6	324	2	D98211	transcription regu
88	43	33.6	413	2	C89780	hypothetical prote
89	43	33.6	427	2	AI2148	processing protein
90	43	33.6	437	2	E71290	probable flagellar
91	43	33.6	441	2	A40707	myosin protein - Ca
92	43	33.6	457	2	S52206	protein-export mem
93	43	33.6	507	2	D75029	hypothetical prote
94	43	33.6	507	2	S74641	sensor histidine k
95	43	33.6	556	2	C87609	hypothetical prote
96	43	33.6	617	2	S27389	secretogranin II -
97	43	33.6	675	2	B56262	exonuclease ABC c
98	43	33.6	675	2	F81095	exonuclease ABC c
99	43	33.6	1128	2	A49960	bud emergence prot
100	43	33.6	1132	2	AD1809	phycobilisome core

## ALIGNMENTS

```

RESULT 1
T18447
HrpW protein - Erwinia amylovora
C:Species: Erwinia amylovora
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T18447
R:Gaudriault, S.
submitted to the EMBL Data Library, May 1998
A:Reference number: Z18936
A:Accession: T18447
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-447 <GAU>
A:Cross-references: UNIPROT:O54508; EMBL:Y13831; NID:e1293450; PID:e1293451; PIDN:CAA741
A:Experimental source: strain CPBP1430; specific host Pommoideae
C:Genetics:
A>Note: hrpW

Query Match 100.0%; Score 128; DB 2; Length 447;
Best Local Similarity 100.0%; Pred. No. 1.9e-10;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSALGQQPIDRTTIEQMAQLLAELLS 27
DB 31 NSALGQQPIDRTTIEQMAQLLAELLS 57

RESULT 2
G69656
2-succinyl-6-hydroxy-2,4-cyclohexadiene-1-carboxylate synthase (EC 4.1.3.-) [validated]
N:Alternate names: menCF protein; menD protein
N:Contains: 2-oxoglutarate decarboxylase (EC 4.1.1.71)
C:Species: Bacillus subtilis
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Aug-2004
C:Accession: G69656; A61649; S27509; T46640; T46641
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertero, C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chelch, E.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabbret, C.; Ferrari, E. Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Gallizzi, A.; Gallier, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Huilo, M.F. Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, A.; Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel, Y. M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle, Rieger, M.; Rivolta, C.; Rocha, E.; Roche, M.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon, A.; Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron, akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama, T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K. T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K. A:Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A>Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A:Reference number: A69580; MUID:98044033; PMID:9384377
A:Accession: G69656
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-580 <KUN>
A:Cross-references: UNIPROT:P23970; GB:299119; GB:AL009126; NID:g2635411; PIDN:CAB15060.
A:Experimental source: strain 168
R:Rowland, B.; Hill, K.; Mueller, J.; Driscoll, J.; Taber, H.
Gene 167, 105-109, 1995
A>Title: Structural organization of a Bacillus subtilis operon encoding menaquinone biosynthesis
A:Reference number: A61649; MUID:96144257; PMID:8566759
A:Accession: A61649
A:Molecule type: DNA
A:Residues: 1-111, 'R', 113-151, 'P', 153-539, 'RRQTS', 545, 'DSI', <RO2>
A:Cross-references: GB:M74538; GB:M74182; GB:M74183; NID:g1185287; PIDN:AAC37014.1; PID:
A>Note: this is a revision to the sequence from reference S27507
R:Rowland, B.; Hill, K.; Mueller, J.; Driscoll, J.; Taber, H.
submitted to the EMBL Data Library, October 1991
A:Description: Organization of an operon involved in menaquinone biosynthesis in Bacillus
A:Reference number: S27507

```

```

A:Accession: S27509
A:Molecule type: DNA
A:Residues: 'MS', 30, 'LKIYT', 36-111, 'R', 113-164, 'ALSM', <ROW1>
A:Cross-references: EMBL:M74538; GB:M74182; GB:M74183; NID:g1185287
A>Note: this sequence has been revised in reference A61649
A:Accession: S27510
A:Molecule type: DNA
A:Residues: 178-539, 'RRQTS', 545, 'DSI', <ROW2>
A:Cross-references: EMBL:M74538; GB:M74182; GB:M74183; NID:g1185287
A>Note: this sequence has been revised in reference A61649
R:Driscoll, J.R.; Taber, H.W. 1992
J. Bacteriol. 174, 5063-5071, 1992
A>Title: Sequence organization and regulation of the Bacillus subtilis menBE operon.
A:Reference number: A42715; MUID:92332443; PMID:1629163
A:Accession: T46640
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 'MS', 30, 'LKIYT', 36-111, 'R', 113-164, 'ALSM', <DRI>
A:Cross-references: EMBL:M74521; NID:g557486; PIDN:AA50398.1; PID:g557489
A:Accession: T46641
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 178-539, 'RRQTS', 545, 'DSI', <DR2>
A:Cross-references: EMBL:M74521; NID:g557486; PIDN:AA50399.1; PID:g557490
C:Genetics:
A:Gene: menD; menCF
A:Start codon: TTG
C:Function:
A:Pathway: ubiquinone biosynthesis
A>Note: thiamine diphosphate cofactor
C:Superfamily: 2-succinyl-6-hydroxy-2,4-cyclohexadiene-1-carboxylate synthase MenD
C:Keywords: carbon-carbon lyase; carboxy-lyase; oxo-acid-lyase

Query Match 39.8%; Score 51; DB 1; Length 580;
Best Local Similarity 40.9%; Pred. No. 23;
Matches 9; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

QY 5 GQQPIDRTTIEQMAQLLAELLK 26
DB 200 GTQSVDRSLSDVAEMLAEEK 221

RESULT 3
T30845
probable DNA repair protein RAD50 - mouse
C:Species: Mus musculus (house mouse)
C>Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 09-Jul-2004
C:Accession: T30845
R:Kim, K.K.; Daud, A.I.; Wong, S.C.; Pajak, L.; Tsai, S.C.; Wang, H.; Henzel, W.J.; Fiecl
J. Biol. Chem. 271, 29255-29264, 1996
A>Title: Mouse RAD50 has limited epitopic homology to p53 and is expressed in the adult n
A:Reference number: Z20899; MUID:97067183; PMID:8910585
A:Accession: T30845
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1312 <KIM>
A:Cross-references: UNIPROT:P70388; EMBL:U66887; NID:g1575574; PID:g1575575; PIDN:AAC5285
C:Genetics:
A:Gene: RAD50
A:Map position: 11
C:Superfamily: RAD50 protein
C:Keywords: DNA repair

Query Match 39.8%; Score 51; DB 2; Length 1312;
Best Local Similarity 47.4%; Pred. No. 55;
Matches 9; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 9 IDRQTIEQMAQLLAELLS 27
DB 1237 LDRENIESLAHVEILKS 1255

RESULT 4

```

AE1803  
sugar binding protein (ABC transporter) homolog lin2972 [imported] - Listeria innocua (C)  
C/Species: Listeria innocua  
C/Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 09-Jul-2004  
C/Accession: AE1803  
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.  
Science 294, 849-852, 2001  
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Maok, C.; Schluster, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend, A.; Title: Comparative genomics of Listeria species.  
A;Reference number: AB1077; MUID:21537279; PMID:11679669  
A;Accession: AE1803  
A;Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-430 <GLA>  
A;Cross-references: UNIPROT:Q926R9; GB:AL592022; PIDN:CAC98197.1; PID:g16415512; GSPDB:C  
A;Experimental source: strain Clip11262  
A;Gene: lin2972

Query Match 39.1%; Score 50; DB 2; Length 430;  
Best Local Similarity 39.1%; Pred.No. 23;  
Matches 9; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

QY 5 GQQPIDRTIEQMQLLAELLKS 27  
| | | : : : : | | | : : :  
Db 341 GAQPVNKEVVEQKAYKENEVIKS 363

RESULT 5  
AF1429  
sugar binding protein (ABC transporter) homolog lmo2839 [imported] - Listeria monocytogenes (C)  
C/Species: Listeria monocytogenes  
C/Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 09-Jul-2004  
C/Accession: AF1429  
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.  
Science 294, 849-852, 2001  
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Maok, C.; Schluster, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend, A.; Title: Comparative genomics of Listeria species.  
A;Reference number: AB1077; MUID:21537279; PMID:11679669  
A;Accession: AF1429  
A;Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-430 <GLA>  
A;Cross-references: UNIPROT:Q8XJG6; GB:NC\_003210; PIDN:CAD01052.1; PID:g16412339; GSPDB:C  
A;Experimental source: strain EGD-e  
C;Genetics:  
A;Gene: lmo2839

Query Match 39.1%; Score 50; DB 2; Length 430;  
Best Local Similarity 39.1%; Pred.No. 23;  
Matches 9; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

QY 5 GQQPIDRTIEQMQLLAELLKS 27  
| | | : : : : | | | : : :  
Db 341 GAQPVNKEVVEQKAYKENEVIKS 363

RESULT 6  
T28155  
variant-specific surface protein 1 - malaria parasite (Plasmodium falciparum) (fragments N;Alternate names: erythrocyte membrane binding protein 1 (EMPI1)  
C/Species: plasmodium falciparum  
C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C/Accession: T28155  
R;Rowe, J.A.; Moulds, J.M.; Newbold, C.I.; Miller, L.H.  
Nature 388, 292-295, 1997  
A;Title: Plasmodium falciparum rosetting is mediated by pFEMP1 and requires complement z

Query Match 38.3%; Score 49; DB 2; Length 902;  
 Best Local Similarity 50.0%; Pred. No. 71;  
 Matches 10; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 5 GQOPDRTQIEQMAQLLAEL 24  
 ||||| : : : : :  
 Db 723 GQOPSDQMTAEQLRDALKNL 742

RESULT 9  
 A96284  
 hypothetical protein AGR\_L\_2452 [imported] - Agrobacterium tumefaciens (strain C58, Cere  
 C:Species: Agrobacterium tumefaciens  
 C>Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 09-Jul-2004  
 C:Accession: A96284  
 R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,  
 A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;  
 Science 294, 2323-2328, 2001  
 A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum  
 A:Reference number: A97359; MUID:21608551; PMID:11743194  
 A:Accession: A96284  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-903 <KUR>  
 A:Cross-References: UNIPROT:Q8U9X3; GB:AE007870; PIDN:AAK89795.1; PID:g15159723; GSPDB:Q  
 C:Genetics:  
 A:Gene: AGR\_L\_2452  
 A:Map position: linear chromosome

Query Match 38.3%; Score 49; DB 2; Length 903;  
 Best Local Similarity 50.0%; Pred. No. 71;  
 Matches 10; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 5 GQOPDRTQIEQMAQLLAEL 24  
 ||||| : : : : :  
 Db 724 GQOPSDQMTAEQLRDALKNL 743

RESULT 10  
 C82953  
 hypothetical protein PA5534 [imported] - Pseudomonas aeruginosa (strain PAO1)  
 C:Species: Pseudomonas aeruginosa  
 C>Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 09-Jul-2004  
 C:Accession: C82953  
 R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Br  
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,  
 .; Lory, S.; Olson, M.V.  
 Nature 406, 959-964, 2000  
 A:Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic patho  
 A:Reference number: A82950; MUID:20437337; PMID:10984043  
 A:Accession: C82953  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-213 <STO>  
 A:Cross-References: UNIPROT:Q9HT40; GB:AE004966; GB:AE004091; NID:g9951872; PIDN:AAG0891  
 A:Experimental source: strain PAO1  
 C:Genetics:  
 A:Gene: PA5534

Query Match 37.9%; Score 48.5; DB 2; Length 213;  
 Best Local Similarity 46.2%; Pred. No. 18;  
 Matches 12; Conservative 6; Mismatches 7; Indels 1; Gaps 1;

QY 2 SALGQOPIDRTQIEQMAQLLAELK 26  
 ||||| : : : : :  
 Db 2 SALPQEKPLPQAFDDDDPQVLAEVLR 27

RESULT 11  
 S76462  
 hypothetical protein - Synecocystis sp. (strain PCC 6803)  
 C:Species: Synecocystis sp.  
 A:Variety: PCC 6803

C>Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 09-Jul-2004  
 C:Accession: S76462  
 R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;  
 O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda,  
 DNA Res. 3, 109-136, 1996  
 A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synecocystis  
 s.  
 A:Reference number: S74322; MUID:97061201; PMID:8905231  
 A:Accession: S76462  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-627 <KAN>  
 A:Cross-References: UNIPROT:P74489; EMBL:D90915; GB:AB001339; NID:g1653604; PIDN:BAA18591  
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 37.9%; Score 48.5; DB 2; Length 627;  
 Best Local Similarity 37.0%; Pred. No. 56;  
 Matches 10; Conservative 8; Mismatches 6; Indels 3; Gaps 1;

QY 2 SALGQOPIDRTQI---EQMAQLLAEL 25  
 : ||||| : : : : :  
 Db 322 TTLGQOPVIEATVVKSQLSQQIVSEVL 348

RESULT 12  
 H83348  
 probable acylase PA2385 [imported] - Pseudomonas aeruginosa (strain PAO1)  
 C:Species: Pseudomonas aeruginosa  
 C>Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 09-Jul-2004  
 C:Accession: H83348  
 R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Br  
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,  
 .; Lory, S.; Olson, M.V.  
 Nature 406, 959-964, 2000  
 A:Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic patho  
 A:Reference number: A82950; MUID:20437337; PMID:10984043  
 A:Accession: H83348  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-762 <STO>  
 A:Cross-References: UNIPROT:Q91194; GB:AE004664; GB:AE004091; NID:g9948415; PIDN:AAG0577:  
 A:Experimental source: strain PAO1  
 C:Genetics:  
 A:Gene: PA2385

Query Match 37.9%; Score 48.5; DB 2; Length 762;  
 Best Local Similarity 32.3%; Pred. No. 69;  
 Matches 10; Conservative 9; Mismatches 3; Indels 9; Gaps 1;

QY 5 GQOPDRTQIEQMA-----QLLAELLK 26  
 ||||| : : : : :  
 Db 523 GKQPLEAKTLEEMVTANHVSADQVLPDLJR 553

RESULT 13  
 S76212  
 hypothetical protein - Synecocystis sp. (strain PCC 6803)  
 C:Species: Synecocystis sp.  
 A:Variety: PCC 6803  
 C>Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 09-Jul-2004  
 C:Accession: S76212  
 R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;  
 O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda,  
 DNA Res. 3, 109-136, 1996  
 A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synecocystis  
 s.  
 A:Reference number: S74322; MUID:97061201; PMID:8905231  
 A:Accession: S76212  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-350 <KAN>  
 A:Cross-References: UNIPROT:P74376; EMBL:D90914; GB:AB001339; NID:g1653477; PIDN:BAA18471  
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 37.5%; Score 48; DB 2; Length 350;  
Best Local Similarity 40.7%; Pred. No. 36;  
Matches 11; Conservative 5; Mismatches 11; Indels 0; Gaps 0;

QY 1 NSALGQOPIRDTIQEQAQLLAELLS 27  
DB 133 NSSIGQIQERQSTSATPQASPELLPS 159

RESULT 14  
D87487  
uridylylate kinase [imported] - Caulobacter crescentus  
C;Species: Caulobacter crescentus  
C;Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 09-Jul-2004  
A;Accession: D87487  
R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.  
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon  
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.  
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
A;Title: Complete Genome Sequence of Caulobacter crescentus.  
A;Reference number: A87249; MUID:21173698; PMID:11259647  
A;Accession: D87487  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-246 <STO>  
A;Cross-references: UNIPROT:Q9A705; GB:AE005673; NID:G1323374; PIDN:AAK23896.1; GSPDB:G  
C;Genetics:  
A;Gene: CCL921  
C;Superfamily: uridylylate kinase

Query Match 36.7%; Score 47; DB 2; Length 246;  
Best Local Similarity 38.5%; Pred. No. 34;  
Matches 10; Conservative 8; Mismatches 6; Indels 2; Gaps 1;

QY 4 LGQOP--IDRQTIEQMAQLLAELLS 27  
DB 23 MGDTPYGDITNTQSVAEADVAVIKS 48

RESULT 15  
S01992  
glutenin low molecular weight chain precursor - wheat  
C;Species: Triticum aestivum (common wheat)  
C;Date: 30-Sep-1989 #sequence\_revision 30-Sep-1989 #text\_change 09-Jul-2004  
A;Accession: S01992  
R;Pitts, E.G.; Rafalaki, J.A.; Hedgcock, C.  
Nucleic Acids Res. 15, 11376, 1988  
A;Title: Nucleotide sequence and encoded amino acid sequence of a genomic gene region fo  
A;Reference number: S01992; MUID:89083577; PMID:3205747  
A;Accession: S01992  
A;Molecule type: DNA  
A;Residues: 1-356 <PIT>  
A;Cross-references: UNIPROT:P10385; EMBL:X07747; NID:G21782; PIDN:CAA30570.1; PID:G21783  
C;Superfamily: gliadin

Query Match 36.7%; Score 47; DB 2; Length 356;  
Best Local Similarity 58.8%; Pred. No. 50;  
Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 4 LGQOPIDRQTIEQMAQL 20  
DB 301 LGQOPQQQLAHQIAQL 317

RESULT 16  
T01866  
hypothetical protein T24M8.3 - Arabidopsis thaliana  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 26-Feb-1999 #sequence\_revision 26-Feb-1999 #text\_change 09-Jul-2004  
A;Accession: T01866  
R;Jataville, P.; Elliott, G.; Le, T.  
submitted to the EMBL Data Library, August 1998

A;Description: The sequence of A. thaliana T24M8.  
A;Reference number: Z14449  
A;Accession: T01866  
A;Status: translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-817 <LAT>  
A;Cross-references: UNIPROT:O81512; EMBL:AF077409; NID:G3319365; PID:G3319369  
A;Experimental source: cultivar Columbia  
C;Genetics:  
A;Map position: 4  
A;Introns: 117/3; 188/2; 269/3; 307/3; 357/3; 395/3; 422/3; 447/1; 486/3; 513/3; 541/3; 5  
A;Note: T24M8.3

Query Match 36.7%; Score 47; DB 2; Length 817;  
Best Local Similarity 44.4%; Pred. No. 1.2e+02;  
Matches 8; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

QY 8 PIDRQTIEQMAQLLAELLS 25  
DB 411 PIDQSSEQQAKIMSELI 428

RESULT 17  
T28130  
hypothetical protein ZK970.6 - Caenorhabditis elegans  
C;Species: Caenorhabditis elegans  
C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
A;Accession: T28130  
R;Berks, M.  
submitted to the EMBL Data Library, April 1995  
A;Reference number: Z20473  
A;Accession: T28130  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-1122 <WIL>  
A;Cross-references: UNIPROT:Q23682; EMBL:Z49073; PIDN:CAA88890.1; GSPDB:GNO0020; CESP:ZK9  
A;Experimental source: clone ZK970  
C;Genetics:  
A;Gene: CESP:ZK970.6  
A;Map position: 2  
A;Introns: 28/3; 72/2; 153/2; 281/1; 312/3; 354/3; 401/1; 442/3; 660/3; 761/1; 819/2; 95  
C;Superfamily: membrane-bound guanylate cyclase; guanylate cyclase catalytic domain homo

Query Match 36.7%; Score 47; DB 2; Length 1122;  
Best Local Similarity 47.4%; Pred. No. 1.7e+02;  
Matches 9; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 6 QOPIRQTIEQMAQLLAEL 24  
DB 798 EQEDRPKAEQICKLUSEM 816

RESULT 18  
PC80  
pancreatic hormone - bovine  
N;Alternate names: pancreatic polypeptide  
C;Species: Bos primigenius taurus (cattle)  
C;Date: 17-Dec-1982 #sequence\_revision 17-Dec-1982 #text\_change 10-May-2001  
A;Accession: A01570  
R;Chance, R.E.; Johnson, M.G.; Hoffmann, J.A.; Jones, W.E.; Koffenberger Jr., J.E.  
unpublished results, cited by Chance, R.E., Moon, N.E., and Johnson, M.G., in Methods of  
San Francisco, and London, 1979  
A;Reference number: A94465  
A;Accession: A01570  
A;Molecule type: protein  
A;Residues: 1-36 <CHA>  
C;Superfamily: pancreatic hormone  
C;Keywords: amidated carboxyl end; hormone; pancreas  
F;1-36/Product: pancreatic hormone #status experimental <PCH>  
F;36/Modified site: amidated carboxyl end (Tyr) #status experimental

Query Match 35.9%; Score 46; DB 1; Length 36;  
Best Local Similarity 57.9%; Pred. No. 6;

Matches 11; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 6 QQPIDRQTIEQMAQLLAEL 24  
: | | | | | | | | | |  
Db 6 EYPGDNATPEQMAQYAAEL 24

RESULT 19  
PCSH  
pancreatic hormone precursor - sheep (tentative sequence) (fragment)  
N;Alternate names: pancreatic polypeptide precursor  
N;Contains: pancreatic hormone; pancreatic icosapeptide  
C;Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)  
C;Date: 25-Feb-1985 #sequence\_revision 25-Feb-1985 #text\_change 09-Jul-2004  
C;Accession: B94465; A91323; A01569  
R;Chance, R.E.; Johnson, M.G.; Hoffmann, J.A.; Jones, W.E.; Koffenberger Jr., J.E.  
unpublished results, cited by Chance, R.E., Moon, N.E., and Johnson, M.G., in Methods of  
San Francisco, and London, 1979  
A;Reference number: A94465  
A;Accession: B94465  
A;Molecule type: protein  
A;Residues: 1-36 <CHA>  
A;Cross-references: UNIPROT:P01301  
R;Schwartz, T.W.; Hansen, H.F.  
FEBS Lett. 168, 293-298, 1984  
A;Title: Isolation of ovine pancreatic icosapeptide: a peptide product containing one cy  
A;Reference number: A91323; PMID:84208777; PMID:6723953  
A;Accession: A91323  
A;Molecule type: protein  
A;Residues: 40-59 <SCH>  
C;Comment: We have added, by homology with the human precursor sequence, Gly-37, which i  
C;Superfamily: pancreatic hormone  
C;Keywords: amidated carboxyl end; hormone; pancreas  
F;1-36/Product: pancreatic hormone #status experimental <PCH>  
F;40-59/Product: pancreatic icosapeptide #status experimental <PCI>  
F;36/Modified site: amidated carboxyl end (Tyr) (amide in mature form from following gly

Query Match 35.9%; Score 46; DB 1; Length 59;  
Best Local Similarity 57.9%; Pred. No. 10;  
Matches 11; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 6 QQPIDRQTIEQMAQLLAEL 24  
: | | | | | | | | | |  
Db 6 EYPGDNATPEQMAQYAAEL 24

RESULT 20  
G91013  
hypothetical protein ECS3079 [imported] - Escherichia coli (strain O157:H7, substrain RI  
C;Species: Escherichia coli  
C;Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 09-Jul-2004  
C;Accession: G91013  
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.  
gasawara, N.; Yasunaga, T.; Kuwara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.  
DNA Res. 8, 11-22, 2001  
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gene  
A;Reference number: A99629; PMID:21156231; PMID:11258796  
A;Accession: G91013  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-75 <HAY>  
A;Cross-references: UNIPROT:P33921; GB:BA000007; PIDN:BA036502.1; PID:g13362548; GSPDB:G  
A;Experimental source: strain O157:H7, substrain RIMD 0509952  
C;Genetics:  
A;Gene: ECS3079  
C;Superfamily: hypothetical protein HI0840

Query Match 35.9%; Score 46; DB 2; Length 75;  
Best Local Similarity 64.7%; Pred. No. 13;  
Matches 11; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 9 IDRQTIEQMAQLLAEL 25  
: | | | | | | | | | |

Search completed: January 25, 2005, 12:15:55  
Job time : 15.9423 secs

Db 4 ISRYSDQEQLLAEL 20

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2005, 12:07:26 ; Search time 101.769 Seconds  
(without alignments)  
152.650 Million cell updates/sec

Title: us-09-879-248-6\_COPY\_31\_57

Perfect score: 128

Sequence: 1 NSALGQPIDRQTIEQMAQLLAEELKLS 27

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : Uniprot 02:\*

1: uniprot\_sprot:\*

2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	128	100.0	138	2 Q79AW7	Q79AW7 erwinia amy
2	128	100.0	447	2 O54508	O54508 erwinia amy
3	128	100.0	447	2 Q9LAW2	Q9LAW2 erwinia amy
4	109	85.2	450	2 Q6XDB8	Q6XDB8 erwinia pyr
5	109	85.2	450	2 AAQ17046	AAQ17046 erwinia p
6	109	85.2	450	2 AAS45453	AAS45453 erwinia p
7	61.5	48.0	1214	2 Q8D6Z1	Q8D6Z1 synchococ
8	53	41.4	416	2 Q6DEJ7	Q6DEJ7 brachydanio
9	53	41.4	440	2 Q6IQ77	Q6IQ77 brachydanio
10	53	41.4	440	2 AAH71533	AAH71533 brachydan
11	53	41.4	943	2 Q7SRG5	Q7SRG5 neurospora
12	52.5	41.0	452	2 Q8U310	Q8U310 pyrococcus
13	51	39.8	377	2 Q7SY40	Q7SY40 brachydanio
14	51	39.8	413	2 Q7NHF2	Q7NHF2 gloebacter
15	51	39.8	498	2 Q8JHT9	Q8JHT9 brachydanio
16	51	39.8	590	1 MEND_BACSU	P23970 b menaquin
17	51	39.8	660	2 Q9CKZ1	Q9CKZ1 pasteurella
18	51	39.8	845	2 Q8BTF5	Q8BTF5 mus musculu
19	51	39.8	911	2 Q8BTF7	Q8BTF7 mus musculu
20	51	39.8	1017	2 Q6VG40	Q6VG40 chimpanzee
21	51	39.8	1017	2 AAR02377	AAR02377 chimpanze
22	51	39.8	1173	2 Q9UP86	Q9UP86 homo sapien
23	51	39.8	1312	2 Q92878	Q92878 homo sapien
24	51	39.8	1312	2 F70388	F70388 mus musculu
25	51	39.8	1312	2 Q9JTL8	Q9JTL8 rattus norv
26	51	39.8	1318	2 Q43254	Q43254 homo sapien
27	50.5	39.5	352	1 LIFO_PSEWI	O05938 pseudomonas
28	50.5	39.5	596	1 ENH_HUMAN	Q96HC4 homo sapien
29	50.5	39.5	596	2 AAT06739	AAT06739 homo sapi
30	50.5	39.5	1263	2 Q7UM20	Q7UM20 rhodopirell
31	50	39.1	430	2 Q926R9	Q926R9 listeria in

32	50	39.1	430	2 Q8Y3J6	Q8Y3J6 listeria mo
33	50	39.1	430	2 Q71VS2	Q71VS2 listeria mo
34	50	39.1	430	2 AAT05594	AAT05594 listeria
35	50	39.1	544	2 Q97473	Q97473 caenorhabdi
36	50	39.1	559	2 Q882U6	Q882U6 pseudomonas
37	50	39.1	2706	2 Q15870	Q15870 plasmodium
38	49.5	38.7	477	2 Q6RJC5	Q6RJC5 candida gla
39	49	38.3	257	2 Q55741	Q55741 synchococyt
40	49	38.3	305	2 Q88Y51	Q88Y51 lactobacill
41	49	38.3	358	2 Q824U8	Q824U8 chlamydomoph
42	49	38.3	457	2 Q7X5J7	Q7X5J7 leuconostoc
43	49	38.3	560	2 Q804Z3	Q804Z3 fugu rubrip
44	49	38.3	628	2 Q841P4	Q841P4 oryza sativ
45	49	38.3	687	2 Q81MA8	Q81MA8 drosophila
46	49	38.3	902	2 Q8U9X3	Q8U9X3 agrobacteri
47	49	38.3	903	2 Q7CSW1	Q7CSW1 giardia lam
48	49	38.3	1016	2 Q7R4T6	Q7R4T6 pseudomonas
49	49	38.3	3470	2 Q88F79	Q88F79 pseudomonas
50	48.5	37.9	213	2 Q9HT40	Q9HT40 pseudomonas
51	48.5	37.9	581	2 Q7VZV8	Q7VZV8 bordetella
52	48.5	37.9	581	2 Q7WC19	Q7WC19 bordetella
53	48.5	37.9	581	2 Q7WQJ5	Q7WQJ5 bordetella
54	48.5	37.9	627	2 P74489	P74489 synchococyt
55	48.5	37.9	762	2 Q91194	Q91194 pseudomonas
56	48	37.5	223	2 Q9V6N0	Q9V6N0 drosophila
57	48	37.5	302	2 Q6ZES1	Q6ZES1 synchococyt
58	48	37.5	302	2 BAD01829	BAD01829 synchococ
59	48	37.5	307	2 Q88SC2	Q88SC2 lactobacill
60	48	37.5	317	2 Q7R711	Q7R711 plasmodium
61	48	37.5	317	2 Q82KJ5	Q82KJ5 salmonella
62	48	37.5	350	2 P74376	P74376 synchococyt
63	48	37.5	441	2 Q88LY0	Q88LY0 pseudomonas
64	48	37.5	463	2 Q8MMJ0	Q8MMJ0 apis cerana
65	48	37.5	491	2 Q8WNE1	Q8WNE1 gorilla gor
66	48	37.5	495	2 Q8WJ2	Q8WJ2 homo sapien
67	48	37.5	522	2 Q7WDI5	Q7WDI5 vibrio vuln
68	48	37.5	522	2 Q8D6M6	Q8D6M6 vibrio vuln
69	48	37.5	584	2 Q700Z9	Q700Z9 pseudomonas
70	48	37.5	584	2 CAD60270	CAD60270 pseudomon
71	48	37.5	1419	1 ESTA_HUMAN	Q86U88 homo sapien
72	48	37.5	4336	2 Q884F8	Q884F8 pseudomonas
73	47.5	37.1	78	2 Q6T8R7	Q6T8R7 ovis aries
74	47.5	37.1	78	2 AAR11453	AAR11453 ovis aries
75	47.5	37.1	131	1 PAHO_BOVIN	F01302 bos taurus
76	47.5	37.1	187	2 Q8DJL3	Q8DJL3 synchococ
77	47.5	37.1	604	2 Q74E16	Q74E16 geobacter s
78	47.5	37.1	604	2 AAR34524	AAR34524 geobacter
79	47	36.7	148	2 Q721R5	Q721R5 thermus the
80	47	36.7	148	2 AAS81409	AAS81409 thermus t
81	47	36.7	202	2 Q8KW98	Q8KW98 ruegeria ep
82	47	36.7	244	2 Q856T2	Q856T2 mycobacteri
83	47	36.7	246	1 PYRH_CAUCR	Q9A705 caulobacter
84	47	36.7	281	2 Q871J0	Q871J0 vibrio para
85	47	36.7	348	2 Q95NK8	Q95NK8 moia macro
86	47	36.7	348	2 Q966U2	Q966U2 moia macro
87	47	36.7	348	2 Q966U3	Q966U3 moia macro
88	47	36.7	356	1 GLTA_WHEAT	F10385 triticum ae
89	47	36.7	357	2 Q6SPY8	Q6SPY8 triticum ae
90	47	36.7	357	2 AAS10192	AAS10192 triticum
91	47	36.7	437	2 Q6W3E3	Q6W3E3 apis mellif
92	47	36.7	437	2 AAQ82184	AAQ82184 apis mell
93	47	36.7	444	2 Q8XH51	Q8XH51 clostridium
94	47	36.7	476	2 Q82880	Q82880 xanthomonas
95	47	36.7	476	2 Q7BTF1	Q7BTF1 xanthomonas
96	47	36.7	476	2 Q921P8	Q921P8 xanthomonas
97	47	36.7	476	2 Q69097	Q69097 xanthomonas
98	47	36.7	487	2 Q9FNR5	Q9FNR5 deutzia rub
99	47	36.7	502	2 Q56831	Q56831 xanthomonas
100	47	36.7	506	2 Q9KH29	Q9KH29 xanthomonas

ALIGNMENTS

```

RESULT 1
Q79AW7 ID Q79AW7 PRELIMINARY; PRT; 138 AA.
AC Q79AW7;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE HrpW (Fragment).
GN Name=hrpW;
OS Erwinia amylovora.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Erwinia.
OX NCBI_TaxID=552;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=Ea321;
RX MEDLINE=98115919; PubMed=9448330;
RA Bogdanove A.J., Kim J.F., Wei Z., Kolchinsky P., Charkowski A.O.,
RA Conlin A.K., Collmer A., Beer S.V.;
RT "Homology and functional similarity of an hrp-linked pathogenicity
RT locus, dspEF, of Erwinia amylovora and the avirulence locus avrE of
RT Pseudomonas syringae pathovar tomato.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:1325-1330(1998).
DR EMBL; U97504; AAC04849.1; -.
FT NON TER 138
SQ SEQUENCE 138 AA; 13788 MW; 4FE177177C74B3C6 CRC64;

Query Match 100.0%; Score 128; DB 2; Length 138;
Best Local Similarity 100.0%; Pred. No. 1.2e-09;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSALGQQPIDRQTIEQMAQLLAELLS 27
Db 31 NSALGQQPIDRQTIEQMAQLLAELLS 57

RESULT 2
O54508 ID O54508 PRELIMINARY; PRT; 447 AA.
AC O54508;
DT 01-NOV-1998 (TrEMBLrel. 06, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE HrpW protein.
GN Name=hrpW;
OS Erwinia amylovora.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Erwinia.
OX NCBI_TaxID=552;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=CFBP1430;
RX MEDLINE=98086111; PubMed=9426142;
RA Gaudriault S., Malandrin L., Paulin J.P., Barny M.A.;
RT "DspA, an essential pathogenicity factor of Erwinia amylovora showing
RT homology with AvrE of Pseudomonas syringae, is secreted via the Hrp
RT secretion pathway in a DspB-dependent way.";
RL Mol. Microbiol. 26:1057-1069(1997).
RN [2]
RN [3]
RN SEQUENCE FROM N.A.
RP STRAIN=CFBP1430;
RX MEDLINE=98316710; PubMed=96541138;
RA Gaudriault S., Briisset M.N., Barny M.A.;
RT "HrpW of Erwinia amylovora, a new Hrp-secreted protein.";
RN [3]
RN SEQUENCE FROM N.A.
RP STRAIN=Ea321;
RX Kim J.F., Zumoff C.H., Beer S.V.;
RA "HrpW, a new harpin of Erwinia amylovora, is a member of a family of
RT pectate lyases.";
RL Phytopathology 87:0-0(1997).

```

```

[4]
RN SEQUENCE FROM N.A.
RP STRAIN=Ea321;
RX MEDLINE=98422475; PubMed=9748455;
RA Kim J.F., Beer S.V.;
RT "HrpW of Erwinia amylovora, a new harpin that contains a domain
RT homologous to pectate lyases of a distinct class.";
RL J. Bacteriol. 180:5203-5210(1998).
DR EMBL; Y13831; CAA74158.1; -.
DR EMBL; U94513; AAC62314.1; -.
DR PIR; T18447; T18447.
DR HSSP; Q9RH00; 1EE6.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0030570; F:pectate lyase activity; IEA.
DR InterPro; IPR011050; Pectin_lyas_like.
DR InterPro; IPR004898; Pect_lyase.
DR Pfam; PF03211; Pectate_lyase; 1.
SQ SEQUENCE 447 AA; 45347 MW; 7A744D169D6D9CF3 CRC64;

Query Match 100.0%; Score 128; DB 2; Length 447;
Best Local Similarity 100.0%; Pred. No. 3.8e-09;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSALGQQPIDRQTIEQMAQLLAELLS 27
Db 31 NSALGQQPIDRQTIEQMAQLLAELLS 57

RESULT 3
Q9LAW2 ID Q9LAW2 PRELIMINARY; PRT; 447 AA.
AC Q9LAW2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Harpin HrpW.
GN Name=hrpW;
OS Erwinia amylovora.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Erwinia.
OX NCBI_TaxID=552;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=Ea246;
RA Kim J.F., Laby R.J., Beer S.V.;
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF083620; AAF63402.1; -.
DR HSSP; Q9RH00; 1EE6.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0030570; F:pectate lyase activity; IEA.
DR InterPro; IPR011050; Pectin_lyas_like.
DR InterPro; IPR004898; Pect_lyase.
DR Pfam; PF03211; Pectate_lyase; 1.
SQ SEQUENCE 447 AA; 45340 MW; 0BBAAE3871EDC2F6 CRC64;

Query Match 100.0%; Score 128; DB 2; Length 447;
Best Local Similarity 100.0%; Pred. No. 3.8e-09;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSALGQQPIDRQTIEQMAQLLAELLS 27
Db 31 NSALGQQPIDRQTIEQMAQLLAELLS 57

RESULT 4
Q6XDB8 ID Q6XDB8 PRELIMINARY; PRT; 450 AA.
AC Q6XDB8;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE HrpW.
GN Name=hrpW;

```



```

OS Erwinia pyrifoliae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Erwinia.
OX NCBI_TaxID=79967;
RN [1]
RC STRAIN=WT#3;
RA Shrestha R., Baek S.J., Bae H.N., Cho J.M., Park D.H., Hwang I.K.,
RA Hur J.H., Lim C.K.;
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
[2]
RN SEQUENCE FROM N.A.
RC STRAIN=WT3;
RA Shrestha R., Baek S.J., Park D.H., Bae H.N., Cho J.M., Hwang I.,
RA Hur J.H., Lim C.K.;
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY237642; AAQ17046.1; -.
DR EMBL; AY530755; AAS45453.1; -.
DR InterPro; IPR011050; Pectin_lyase_like.
DR Pfam; PF03211; Pectate_lyase; 1.
SQ SEQUENCE 450 AA; 45509 MW; 085F90CD322C3A74 CRC64;

Query Match      85.2%; Score 109; DB 2; Length 450;
Best Local Similarity 88.5%; Pred. No. 1.6e-06;
Matches 23; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 NSALGQQPIDRQTIEQMAQLLAELLK 26
Db 31 NSAQGHQPIDRQTIEQMAQLLGELLK 56

RESULT 5
AAQ17046
ID AAQ17046 PRELIMINARY; PRT; 450 AA.
AC AAQ17046;
DT 03-MAR-2004 (T-EMBLrel. 27, Created)
DT 03-MAR-2004 (T-EMBLrel. 27, Last sequence update)
DE HrpW.
GN Erwinia pyrifoliae.
OS Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Erwinia.
OX NCBI_TaxID=79967;
RN [1]
RC STRAIN=WT#3;
RA Shrestha R., Baek S.J., Bae H.N., Cho J.M., Park D.H., Hwang I.K.,
RA Hur J.H., Lim C.K.;
RL "Molecular characterization hrp genes cluster of Erwinia pyrifoliae
RT and expression of hrpB encoding elicitor of the hypersensitive
RT response."
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY237642; AAQ17046.1; -.
SQ SEQUENCE 450 AA; 45509 MW; 085F90CD322C3A74 CRC64;

Query Match      85.2%; Score 109; DB 2; Length 450;
Best Local Similarity 88.5%; Pred. No. 1.6e-06;
Matches 23; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 NSALGQQPIDRQTIEQMAQLLAELLK 26
Db 31 NSAQGHQPIDRQTIEQMAQLLGELLK 56

RESULT 6
AAS45453
ID AAS45453 PRELIMINARY; PRT; 450 AA.
AC AAS45453;
DT 02-MAR-2004 (T-EMBLrel. 27, Created)
DT 02-MAR-2004 (T-EMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (T-EMBLrel. 27, Last annotation update)

```

```

DE HrpW.
GN HRPW.
OS Erwinia pyrifoliae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Erwinia.
OX NCBI_TaxID=79967;
RN [1]
RC STRAIN=WT#3;
RA Shrestha R., Baek S.J., Park D.H., Bae H.N., Cho J.M., Hwang I.,
RA Hur J.H., Lim C.K.;
RT "Identification of hrp genes cluster and characterization of HR
RT elicitor hrpNEP gene in Erwinia pyrifoliae."
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY530755; AAS45453.1; -.
SQ SEQUENCE 450 AA; 45509 MW; 085F90CD322C3A74 CRC64;

Query Match      85.2%; Score 109; DB 2; Length 450;
Best Local Similarity 88.5%; Pred. No. 1.6e-06;
Matches 23; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 NSALGQQPIDRQTIEQMAQLLAELLK 26
Db 31 NSAQGHQPIDRQTIEQMAQLLGELLK 56

RESULT 7
Q8DGZ1
ID Q8DGZ1 PRELIMINARY; PRT; 1214 AA.
AC Q8DGZ1;
DT 01-MAR-2003 (T-EMBLrel. 23, Created)
DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
DE Hydantoinease / oxoprolinase.
GN OrderedLocusNames=tlr2171;
OS Synechococcus elongatus (Thermosynechococcus elongatus).
OC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.
OX NCBI_TaxID=32046;
RN [1]
RC STRAIN=BP-1;
RC STRAIN=BP-1;
RX MEDLINE=22225144; PubMed=12240834;
RA Nakamura Y., Kaneko T., Sato S., Ikeuchi M., Katoh H., Sasamoto S.,
RA Watanabe A., Iriguchi M., Kawashima K., Kimura T., Kishida Y.,
RA Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Nakazaki N.,
RA Shimo S., Sugimoto M., Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the thermophilic cyanobacterium
RT Thermosynechococcus elongatus BP-1."
RL DNA Res. 9:123-130(2002).
DR EMBL; AP005376; BAC09723.1; -.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR InterPro; IPR002821; Hydantoinease_A.
DR InterPro; IPR003692; Hydantoinease_B.
DR InterPro; IPR008040; Hydant_A_N.
DR Pfam; PF01968; Hydantoinease_A; 1.
DR Pfam; PF02538; Hydantoinease_B; 1.
DR Pfam; PF05378; Hydant_A_N; 1.
KW Complete proteome.
SQ SEQUENCE 1214 AA; 129918 MW; 929531C290F24447 CRC64;

Query Match      48.0%; Score 61.5; DB 2; Length 1214;
Best Local Similarity 54.2%; Pred. No. 15;
Matches 13; Conservative 5; Mismatches 5; Indels 1; Gaps 1;

QY 5 GQPIDRQTIE-QMAQLLAELLK 27
Db 395 GQPLDRDSVEQPAQLRQEIYRS 418

RESULT 8
Q6DEJ7
ID Q6DEJ7 PRELIMINARY; PRT; 416 AA.
AC Q6DEJ7;

```

01-OCT-2004	(TRENBLrel. 28, Created)
01-OCT-2004	(TRENBLrel. 28, Last sequence update)
01-OCT-2004	(TRENBLrel. 28, Last annotation update)
Hypothetical protein.	
Brachydanio rerio (Zebrafish)	(Danio rerio).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;	
Cyprinidae; Danio.	
NCBI_TaxID=7955;	
[1]	
SEQUENCE FROM N.A.	
STRAIN=Singapore local strain; TISSUE=Embryo;	
PubMed=12477932;	
Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,	
Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,	
Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,	
Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,	
Diatchenko L., Marusina K., Farmer A.S., Rubin G.M., Hong L.,	
Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,	
Raha S.S., Loquellano N.A., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,	
Bosak S.A., McEwan P.J., Peters G.J., Abramson R.D., Mullaby S.J.,	
Richards S., Worley K.C., McKernan K.J., Malek J.A., Gunaratne P.H.,	
Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,	
Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,	
Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,	
Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,	
Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,	
Krzywinski M.I., Skalek U., Smailus D.E., Schnerch A., Schein J.E.,	
Jones S.J., Marra M.A.	
"Generation and initial analysis of more than 15,000 full-length human	
and mouse cDNA sequences."	
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).	
[2]	
SEQUENCE FROM N.A.	
STRAIN=Singapore local strain; TISSUE=Embryo;	
Strausberg R.	
Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.	
EMBL; BC071115; AAH77115.1; -.	
Hypothetical protein.	
SEQUENCE 416 AA; 47726 MW; 008C38F5B606CE42 CRC64;	
Query Match 41.4%; Score 53; DB 2; Length 416;	
Best Local Similarity 38.5%; Pred. No. 79;	
Matches 10; Conservative 8; Mismatches 8; Indels 0; Gaps 0;	
Qy 1 NSALGOQPIDRQTIEQMAQLAELLK 26	
: : :	
Db 57 NPLGKKGIESRVLEESQYLIAEMLK 82	
RESULT 9	
Q61Q77	
ID Q61Q77 PRELIMINARY; PRT; 440 AA.	
AC Q61Q77;	
DT 05-JUL-2004 (TRENBLrel. 27, Created)	
DT 05-JUL-2004 (TRENBLrel. 27, Last sequence update)	
DT 05-JUL-2004 (TRENBLrel. 27, Last annotation update)	
DE Zgc:86915.	
GN Name=zgc:86915;	
OS Brachydanio rerio (Zebrafish) (Danio rerio).	
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;	
Cyprinidae; Danio.	
NCBI_TaxID=7955;	
[1]	
SEQUENCE FROM N.A.	
STRAIN=Singapore local strain; TISSUE=Embryo;	
PubMed=12477932;	
Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,	
Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,	
Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,	
Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,	

RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RL and mouse cDNA sequences.";  
 RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Embryo;  
 RA Strausberg R.;  
 RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.  
 KW EMBL; BC071533; AAH71533.1; -;  
 SQ SEQUENCE 440 AA; 50312 MW; D1D4740B11EFF193 CRC64;  
 Query Match 41.4%; Score 53; DB 2; Length 440;  
 Best Local Similarity 42.3%; Pred. No. 84;  
 Matches 11; Conservative 6; Mismatches 9; Indels 0; Gaps 0;  
 QY 1 NSALGQPIDRQIEQMAQLLAELLK 26  
 DB 81 NFGLGKKSIESRVLESQYLFALLK 106  
 RESULT 11  
 Q7SBG5 PRELIMINARY; PRT; 943 AA.  
 ID AC Q7SBG5  
 DT 01-MAR-2004 (T-EMBLrel. 26, Created)  
 DT 01-MAR-2004 (T-EMBLrel. 26, Last sequence update)  
 DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)  
 DE Hypothetical protein.  
 GN Name=NCU07591.1;  
 OS Neurospora crassa.  
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
 OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.  
 OX NCBI\_TaxID=5141;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=OR74A;  
 RA Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D.,  
 RA Jaffe D., Fitzhugh W., Ma L.-J., Smirnov S., Purcell S., Rehman B.,  
 RA Elkins T., Engels R., Wang S., Nielsen C.B., Butler J., Endrizzi M.,  
 RA Qui D., Iankiev P., Pedersen D., Nelson M., Washburne M.,  
 RA Selitrennikoff C.P., Kinsey J.A., Braun E.L., Zelter A., Schulte U.,  
 RA Kothe G.O., Jedd G., Mewes W., Staben C., Marcotte E., Greenberg D.,  
 RA Roy A., Foley K., Naylor J., Thomann N., Barrett R., Gnerre S.,  
 RA Kamal M., Kamveselis M., Mauceli E., Bieleke C., Rudd S., Frishman D.,  
 RA Kryzofova S., Rasmussen C., Metznerberg R.L., Perkins D.D., Kroken S.,  
 RA Cogoni C., Macino G., Catcheside D., Li W., Pratt R.J., Osmani S.A.,  
 RA Desouza C.C., Glass L., Orbach M.J., Berglund J., Voelker R.,  
 RA Yarden O., Plamann M., Seiler S., Dunlap J., Radford A., Aramayo R.,  
 RA Natvig D.O., Alex L.A., Mannhaupt G., Ebbole D.J., Freitag M.,  
 RA Paulsen I., Sachs M.S., Lander E.S., Nusbaum C., Birren B.,  
 RT "The Genome Sequence of the Filamentous Fungus Neurospora crassa.";  
 RL Nature 0:0-0(2003).  
 CC -1- CAUTION: The sequence shown here is derived from an  
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
 CC preliminary data.  
 DR EMBL; AABX01000168; EAA33735.1; -;  
 KW Hypothetical protein.  
 SQ SEQUENCE 943 AA; 105278 MW; 53586D8C6BDD7AF6 CRC64;  
 Query Match 41.4%; Score 53; DB 2; Length 943;  
 Best Local Similarity 40.9%; Pred. No. 1.8e+02;  
 Matches 9; Conservative 8; Mismatches 5; Indels 0; Gaps 0;  
 QY 4 LGQPIDRQIEQMAQLLAEL 25  
 DB 536 VGGQPLDRVLQTLSQLREIL 557  
 RESULT 12  
 Q8U310 PRELIMINARY; PRT; 452 AA.  
 ID AC Q8U310  
 DT 01-JUN-2002 (T-EMBLrel. 21, Created)  
 DT 01-JUN-2002 (T-EMBLrel. 21, Last sequence update)  
 DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)  
 DE Putative nucleolar protein IV (Nol1-nop2-sun family).  
 GN OrderedLocusNames=PF0666;  
 OS Pyrococcus furiosus.  
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;  
 OC Pyrococcus.  
 OX NCBI\_TaxID=2261;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=DSM 3638;  
 RX MEDLINE=21079003; PubMed=11210495;  
 RA Robb F.T., Maeder D.L., Brown J.R., Diruggiero J., Stump M.D.,  
 RA Yeh R.K., Weiss R.B., Dunn D.M.;  
 RT "Genomic sequence of hyperthermophile, Pyrococcus furiosus:  
 RT implications for physiology and enzymology.";  
 RL Meth. Enzymol. 330:134-157(2001).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Vc1 / DSM 3638 / ATCC 43587 / JCM 8422;  
 RA Weiss R.B., Dunn D.M., Robb F.T., Brown J.R.;  
 RT "The complete sequence of the Pyrococcus furiosus genome.";  
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AE010187; AAL80790.1; -;  
 DR GO; GO:0003723; F:RNA binding; IEA.  
 DR GO; GO:0008757; F:S-adenosylmethionine-dependent methyltransf.; IEA.  
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
 DR InterPro; IPR001678; Fmu\_NOL1/Nop2p.  
 DR InterPro; IPR011023; Nop2p.  
 DR InterPro; IPR006027; NusB.  
 DR InterPro; IPR006174; NusB region.  
 DR InterPro; IPR000051; SAM bind.  
 DR Pfam; PF01189; NOL1\_Nop2\_Fmu; 1.  
 DR ProDom; PD005242; NusB\_region; 1.  
 DR TIGRFAMs; TIGR00446; nop2p; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 452 AA; 52212 MW; 99FFDD7419D4583E CRC64;  
 Query Match 41.0%; Score 52.5; DB 2; Length 452;  
 Best Local Similarity 40.8%; Pred. No. 1e+02;  
 Matches 13; Conservative 6; Mismatches 8; Indels 5; Gaps 1;  
 QY 1 NSALGQGP-----IDRQIEQMAQLLAELK 27  
 DB 344 SGTIGKPNELRWLRQKIEEWAQLQRELLS 375  
 RESULT 13  
 Q7SY40 PRELIMINARY; PRT; 377 AA.  
 ID AC Q7SY40  
 DT 01-OCT-2003 (T-EMBLrel. 25, Created)  
 DT 01-OCT-2003 (T-EMBLrel. 25, Last sequence update)  
 DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)  
 DE Zgc:63546.  
 GN Name=zgc:63546;  
 OS Brachydanio rerio (Zebrafish) (Danio rerio).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Cypriniformes;  
 OC Cyprinidae; Danio.  
 OX NCBI\_TaxID=7955;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=AB; TISSUE=whole body;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heish F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,



RT "Sequence organization and regulation of the *Bacillus subtilis* menBE  
operon";  
RL J. Bacteriol. 174:5063-5071(1992).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=168 / RBI;  
RX MEDLINE=96144257; PubMed=8566759;  
RA Rowland B., Hill K., Miller P., Driscoll J.R., Taber H.W.;  
RT "Structural organization of a *Bacillus subtilis* operon encoding  
menaquinone biosynthetic enzymes.";  
RL Gene 167:105-109(1995).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=168;  
RX MEDLINE=98048467; PubMed=9387221;  
RA Lapidus A., Galleron N., Sorokin A., Ehrlich S.D.;  
RT "Sequencing and functional annotation of the *Bacillus subtilis* genes  
in the 200 kb rnbB-dnaB region";  
RL Microbiology 143:3431-3441(1997).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN=168;  
RX MEDLINE=98044033; PubMed=9384377; DOI=10.1038/36786;  
RA Kunst F., Ogawara N., Moszer I., Albertini A.M., Alloni G.,  
Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,  
Borries R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,  
Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,  
Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,  
Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,  
Entian K.-D., Errington J., Fabret C., Ferrari E., Foulger D.,  
Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,  
Ghim S.Y., Glaeser P., Goffeau A., Golightly E.J., Grandi G.,  
Guiseppi G., Guy B.J., Haga K., Hatesch J., Harwood C.R., Henaut A.,  
Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M.,  
Jones L.-M., Joris B., Karamata D., Kasahara Y., Kleer-Blanchard M.,  
Klein C., Kobayashi Y., Koetter P., Koningsstein G., Krogh S.,  
Kumano M., Kurita K., Lapidus A., Lardinois S., Lauber J.,  
Lazarevic V., Lee S.M., Levine A., Liu H., Masuda S., Mauel C.,  
Medigue C., Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S.,  
Noback M., Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B.,  
Park S.H., Parro V., Pohl T.M., Portetelle D., Porwollik S.,  
Prescott A.M., Presecan E., Pujic P., Purnelle B., Rappoport G.,  
Rey M., Reynolds S., Rieger M., Rivolta C., Rocha E., Roche B.,  
Rose M., Sadaie Y., Sato T., Scanlan E., Schleich S., Schroeter R.,  
Scoffone F., Sekiguchi J., Sekowska A., Seror S.J., Seror P.,  
Shin B.S., Soldo B., Sorokin A., Tacconi E., Takagi T., Takahashi H.,  
Takenaru K., Takeuchi M., Tanakoshi A., Tanaka T., Terpetra P.,  
Tognoni A., Tosato V., Uchiyama S., Vandenbol M., Vannier P.,  
Vassarotti A., Viari A., Wambutt R., Wedler E., Wedler H.,  
Weissenegger T., Winters P., Wipat A., Yamamoto H., Yamane K.,  
Yasunoto K., Yata K., Yoshida K., Yoshikawa H.F., Zumbstein E.,  
Yoshikawa H., Danchin A.;  
RT "The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*";  
RL Nature 390:249-256(1997).  
CC -1- CATALYTIC ACTIVITY: isochlorismate + 2-ketoglutarate = 2-succinyl-  
6-hydroxy-2,4-cyclohexadiene-1-carboxylate + pyruvate + CO(2).  
CC -1- CATALYTIC ACTIVITY: 2-oxoglutarate = succinate semialdehyde +  
CO(2).  
CC -1- CATALYTIC ACTIVITY: 2-oxoglutarate + isochlorismate = (1S,6R)-6-  
hydroxy-2-succinylcyclohexa-2,4-diene-1-carboxylate + pyruvate +  
CO(2).  
CC -1- COPACTOR: Binds 1 thiamine pyrophosphate and 1 magnesium ion per  
subunit (By similarity).  
CC -1- PATHWAY: Menaquinone biosynthesis; first step.  
CC -1- SIMILARITY: Belongs to the TPP enzyme family.  
CC -1- CAUTION: Used to include what was called 'menCF'.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; M74521; AAA50398.1; ALT\_SEQ.  
DR EMBL; M74521; AAA50399.1; ALT\_SEQ.  
DR EMBL; M74538; AAC37014.1; -.  
DR EMBL; AF008220; AAC00224.1; -.  
DR EMBL; Z99119; CAB15060.1; -.  
DR PIR; G69656; G69656.  
DR Subtilisin; BG10683; menD.  
DR TIGRFA; TIGR00173; menD; 1.  
DR Complete proteome; Flavoprotein; Lyase; Menaquinone biosynthesis;  
KW Multifunctional enzyme; Thiamine pyrophosphate; Transferase.  
FT CONFLICT 112 112 E -> R (in Ref. 2).  
FT CONFLICT 152 152 A -> P (in Ref. 2).  
FT CONFLICT 540 580 PQADKPLHLIETKDRQSRVQLHRDMLNEAVREVKKQWEL  
FT CONFLICT 540 580 -> RROTSPDSI (in Ref. 2).  
SQ SEQUENCE 580 AA; 64092 MW; 3B416F8DA18FFAF2 CRC64;  
  
Query Match 39.8%; Score 51; DB 1; Length 580;  
Best Local Similarity 40.9%; Pred. NO. 2.1e+02;  
Matches 9; Conservative 7; Mismatches 6; Indels 0; Gaps 0;  
  
QY 5 GQPIDRQTIEQMAQLAEELK 26  
DB 200 GTQSDRESLSDDVAEMLAEEK 221  
  
RESULT 17  
Q9CKZ1 PRELIMINARY; PRT; 660 AA.  
ID Q9CKZ1  
AC Q9CKZ1;  
DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE PgtB.  
GN Name:pgtB; OrderedLocusNames=PM1459;  
OS Pasteurella multocida.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;  
OC Pasteurellaceae; Pasteurella.  
OX NCBI\_TaxID=747;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Pm70;  
RX MEDLINE=21145866; PubMed=11248100;  
RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;  
RT "Complete genomic sequence of *Pasteurella multocida* Pm70";  
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).  
CC -1- SIMILARITY: Contains 1 histidine kinase domain.  
DR EMBL; AB006182; AAK03543.1; -.  
DR GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:0005524; F:ATP binding; IEA.  
DR GO; GO:0016301; F:kinase activity; IEA.  
DR GO; GO:0000155; F:cis-acting molecule activity; IEA.  
DR GO; GO:0007165; P:signal transduction; IEA.  
DR InterPro; IPR003594; ATPbind\_ATPase.  
DR InterPro; IPR003660; HAMF.  
DR InterPro; IPR005467; His\_kinase.  
DR InterPro; IPR003661; His\_kin\_N.  
DR InterPro; IPR009082; His\_kin\_Homodim.  
DR Pfam; PF00672; HAMF; 1.  
DR Pfam; PF02518; HAMF; 1.  
DR Pfam; PF00512; HSKA; 1.  
DR SMART; SM00304; HAMF; 1.  
DR SMART; SM00387; HAMF; 1.  
DR SMART; SM00388; HSKA; 1.  
DR PROSITE; PS00885; HAMF; 1.  
DR PROSITE; PS0109; HIS\_KIN; 1.  
KW Complete proteome; Kinase; Phosphorylation; Sensory transduction;  
KW Transferase.  
SQ SEQUENCE 660 AA; 75289 MW; E461F1C11FFE003A CRC64;

Query Match 39.8%; Score 51; DB 2; Length 660;  
 Best Local Similarity 55.0%; Pred. No. 2.4e+02;  
 Matches 11; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

OY 6 QQPIDRQTIEOMALAEEL 25  
 DB 178 QQPNNKQIEQLKQOELL 197  
 ||||:|||||:|

RESULT 18  
 Q8BTF5 ID Q8BTF5 PRELIMINARY; PRT; 845 AA.  
 AC Q8BTF5  
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Mus musculus 2 days neonate thymus thymic cells cDNA, RIKEN full-length enriched library, clone:C920014N24 product:DNA LIGASE IV (EC 6.5.1.1) (POLYDEOXYRIBONUCLEOTIDE SYNTHASE).  
 DE 6.5.1.1 (POLYDEOXYRIBONUCLEOTIDE SYNTHASE).  
 GN Name-Lig4;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxId=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Thymus;  
 RX MEDLINE=99279253; PubMed=10349636;  
 RA Carninci P., Hayashizaki Y.;  
 RT "High-efficiency full-length cDNA cloning.";  
 RL Meth. Enzymol. 303:19-44(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Thymus;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA RIKEN FANTOM Consortium;  
 RT "Functional annotation of a full-length mouse cDNA collection.";  
 RL Nature 409:685-690(2001).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Thymus;  
 RX MEDLINE=20499374; PubMed=11042159;  
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M., Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;  
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes.";  
 RL Genome Res. 10:1617-1630(2000).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Thymus;  
 RX MEDLINE=20499374; PubMed=11042159;  
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M., Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;  
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes.";  
 RL Genome Res. 10:1617-1630(2000).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Thymus;  
 RX MEDLINE=20530913; PubMed=11076861;  
 RA Shibata K., Itoh M., Aizawa K., Nagaoaka S., Sasaki N., Carninci P., Konno H., Akiyama J., Nishi K., Kiteunai T., Tashiro H., Itoh M., Sumi N., Ishii Y., Nakamura S., Hazama M., Nishino T., Harada A., Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K., Fujiwaki S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M., Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J., Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;  
 RT "RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer.";  
 RL Genome Res. 10:1757-1771(2000).  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Thymus;  
 RX MEDLINE=99279253; PubMed=10349636;  
 RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,

RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W., Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T., Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T., Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S., Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M., Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Sano H., Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H., Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M., Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T., Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;  
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
 CC -1- CATALYTIC ACTIVITY: ATP + [deoxyribonucleotide] (N) + [deoxyribonucleotide] (M) = AMP + diphosphate + [deoxyribonucleotide] (N+M).  
 CC -1- SIMILARITY: Belongs to the ATP-dependent DNA ligase family.  
 CC EMBL; AK083353; BAC38882.1; -.  
 DR HSP; P49917; I1K9.  
 DR MGD; MGI:1335098; Lig4.  
 DR GO; GO:0005622; C:intracellular; IEA.  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0003910; F:DNA ligase (ATP) activity; IEA.  
 DR GO; GO:0016874; F:ligase activity; IEA.  
 DR GO; GO:0003676; F:nucleic acid binding; IEA.  
 DR GO; GO:0006310; P:DNA recombination; IEA.  
 DR GO; GO:0006281; P:DNA repair; IEA.  
 DR GO; GO:0006260; P:DNA replication; IEA.  
 DR InterPro; IPR001357; BRCT.  
 DR InterPro; IPR000977; DNA\_ligase.  
 DR InterPro; IPR008994; Nucleic\_acid\_OB.  
 DR Pfam; PF00533; BRCT; 2.  
 DR Pfam; PF04679; DNA\_ligase\_A\_C; 1.  
 DR Pfam; PF01068; DNA\_ligase\_A\_M; 1.  
 DR Pfam; PF04675; DNA\_ligase\_A\_N; 1.  
 DR SMART; SM00292; BRCT; 2.  
 DR TIGRFAM; TIGR00574; dnll; 1.  
 DR PROSITE; PS00172; BRCT; 2.  
 DR PROSITE; PS00697; DNA\_LIGASE\_A1; UNKNOWN\_1.  
 DR PROSITE; PS00333; DNA\_LIGASE\_A2; 1.  
 DR PROSITE; PS00160; DNA\_LIGASE\_A3; 1.  
 KW ATP-binding; DNA recombination; DNA repair; DNA replication; Ligase.  
 SQ SEQUENCE 845 AA; 96487 MW; 9CA8A6B935C259C CRC64;

Query Match 39.8%; Score 51; DB 2; Length 845;  
 Best Local Similarity 47.6%; Pred. No. 3e+02;  
 Matches 10; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

OY 4 LGQPIDRQTIEOMALAEEL 24  
 DB 713 LGIRPSSQQTPEWAPVIADL 733  
 ||||:|||||:|

RESULT 19  
 Q8BTF7 ID Q8BTF7 PRELIMINARY; PRT; 911 AA.  
 AC Q8BTF7  
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Mus musculus adult male testis cDNA, RIKEN full-length enriched library, clone:4932416F16 product:DNA LIGASE IV (EC 6.5.1.1) (POLYDEOXYRIBONUCLEOTIDE SYNTHASE).  
 GN Name-Lig4;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxId=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Testis;  
 RX MEDLINE=99279253; PubMed=10349636;  
 RA Carninci P., Hayashizaki Y.;  
 RT "High-efficiency full-length cDNA cloning.";  
 RL Meth. Enzymol. 303:19-44(1999).



Thu Jan 27 12:40:01 2005

us-09-879-248-6\_copy\_31\_57.rup

Page 10

Db 184 LKQWPLSRKIEALTQITQEMLK 206

Search completed: January 25, 2005, 12:15:27  
Job time : 111.769 secs